

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

protein - protein search, using sw model

on: February 25, 2004, 14:37:48 ; Search time 55 Seconds  
(without alignments)  
107.882 Million cell updates/sec

le: US-09-915-914B-7

fect score: 135

uence: 1 KETWETWTWTSOPKKRKV 21

ring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 1586107 seqs, 282547505 residues

al number of hits satisfying chosen parameters: 1586107

imum DB seq length: 0

imum DB seq length: 2000000000

t-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase :

A: Genesep23Jan04:\*

1: Genesep1980s:\*

2: Genesep1990s:\*

3: Genesep2000s:\*

4: Genesep2001s:\*

5: Genesep2002s:\*

6: Genesep2003as:\*

7: Genesep2003bs:\*

8: Genesep2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

ult No.	Score	Query Match	Length	DB ID	Description
1	135	100.0	21	5	ABG78995 Cell pene
2	135	100.0	21	5	ABG78995 Peptide t
3	135	100.0	21	5	ABG78995 Peptide t
4	135	100.0	21	7	ABG78995 Peptide s
5	135	100.0	21	7	ABG78995 Protein-d
6	135	91.1	21	5	ABG78995 Peptide t
7	120	88.9	20	5	ABG78995 Peptide t
8	108	80.0	19	5	ABG78995 Peptide t
9	106	78.5	20	5	ABG78995 Peptide t
10	105.5	78.1	20	5	ABG78995 Peptide t
11	102	75.6	19	5	ABG78995 Peptide t
12	99	73.7	19	5	ABG78995 Peptide t
13	99	66.3	12	5	ABG78995 Hydropob
14	90	66.7	12	5	ABG78995 Cell-targ
15	90	66.7	21	5	ABG78995 Generic p
16	90	66.7	26	5	ABG78995 Chimeric
17	85	63.0	19	5	ABG78995 Peptide t
18	79	58.5	11	5	ABG78995 Peptide t
19	73	54.1	17	5	ABG78995 Peptide t
20	69	51.1	9	5	ABG78995 Peptide t
21	69	51.1	18	5	ABG78995 Peptide t
22	69	51.1	19	5	ABG78995 Peptide t
23	69	51.1	19	5	ABG78995 Peptide t
24	69	51.1	19	5	ABG78995 Peptide t
25	69	51.1	20	5	ABG78995 Generic p

26	64	47.4	19	5	ABG78995
27	61.5	45.6	430	5	ABG78995
28	61.5	45.6	560	4	ABG78995
29	61.5	45.6	560	6	ABG78995
30	61.5	45.6	560	7	ABG78995
31	61.5	45.6	562	6	ABG78995
32	61.5	45.6	562	6	ABG78995
33	61.5	45.6	565	6	ABG78995
34	61.5	45.6	724	5	ABG78995
35	61.5	45.6	739	5	ABG78995
36	61.5	45.6	760	1	ABG78995
37	61.5	45.6	850	4	ABG78995
38	61.5	45.6	850	4	ABG78995
39	61.5	45.6	850	5	ABG78995
40	61.5	45.6	850	6	ABG78995
41	61.5	45.6	850	6	ABG78995
42	61.5	45.6	850	6	ABG78995
43	61.5	45.6	875	4	ABG78995
44	61.5	45.6	875	4	ABG78995
45	61.5	45.6	875	5	ABG78995

## ALIGNMENTS

### RESULT 1

ABG78995  
ID ABG78995 standard; peptide; 21 AA.

XX AC ABG78995;

XX DT 15-NOV-2002 (first entry)

XX DE Cell penetrating peptide, peptide carrier.

XX KW Cell penetrating peptide; cancer; tumour; melanoma; lymphoma; sarcoma; lung cancer; non-Hodgkin's lymphoma; leukaemia; Hodgkin's lymphoma; uterine cancer; cervical cancer; bladder cancer; kidney cancer; adenocarcinoma; breast cancer; prostate cancer; ovarian cancer; pancreatic cancer; vaccine; dendritic cell; tumour infiltrating lymphocyte; TIL; human leukocyte antigen; HLA; cytostatic.

XX OS Unidentified.

XX PN WO200264057-A2.

XX PD 22-AUG-2002.

XX PF 15-FEB-2002; 2002WO-US005212.

XX PR 15-FEB-2001; 2001US-0268687P.

XX PA (BAYU) BAYLOR COLLEGE MEDICINE.

XX PI Wang R;

XX PT WPI; 2002-627577/67.

XX DR Novel composition for treating a disease in an animal, comprises an immune effector cell and cell penetrating peptide associated with an antigen or antibody.

XX PS Disclosure; Page 11; 61pp; English.

XX CC The invention relates to a composition (I) comprising an immune effector cell and a cell penetrating peptide (CPP) associated with an antigen or antibody. Also included are (1) a vaccine comprising (I), CPP associated with an antigen, and a pharmaceutically acceptable carrier and (2) preparing a composition for a disease, by providing (I) and CPP associated with an antigen for disease, and introducing the antigen-associated CPP to (I), where antigen enters into the cell. The antigens are, for example, tumour antigen derived epitopes recognised by tumour

Example 4; Page 77; 156pp; English.

The invention relates to a transfection agent comprises a peptide of about 16 - 30 amino acids in length. Peptides of the invention comprise a hydrophobic domain, a hydrophilic domain, optionally a spacer sequence between the domains and a functional group conjugated to at least one terminal of the peptide. Peptides of the invention are useful for a non-covalent association with and transport of a heterologous compound into a cell. They are also useful for promoting the cellular internalisation of at least one member e.g. peptide, proteins, antibodies, their derivatives

The invention relates to a transfection agent comprising a peptide of about 16 - 30 amino acids in length. Peptides of the invention comprise a hydrophobic domain, a hydrophilic domain, optionally a spacer sequence between the domains and a functional group conjugated to at least one terminal of the peptide. Peptides of the invention are useful for a non-covalent association with and transport of a heterologous compound into a cell. They are also useful for promoting the cellular internalisation of at least one member e.g. peptide, proteins, antibodies, their derivative and/or conjugates. They may form part of a pharmaceutical composition to deliver the compound selected from a diagnostic or therapeutic compound,

to treat at least one condition such as cancer or an infectious disease, or which targets a cancerous cell or pathogen-infected cell and to deliver a peptide or inhibitor that disrupts the activity of the enzyme. The agent of the invention has a transfection efficiency of at least 5% for at least two of the members of the group of the compounds. The agent has a good delivery efficiency for a broad spectrum of compounds and cell types, has a low toxicity, are easy to handle and easy to formulate in conjunction with the many different compound types that it can deliver. The peptides are serum sensitive, thus they bode particularly well for systemic and/or localised in patients. The current sequence represents the peptide transfection agent Pep-2

Sequence 21 AA;

Query Match 100.0%; Score 135; DB 5; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.3e-10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KETWETWETWSQPKKRV 21  
|||||  
1 KETWETWETWSQPKKRV 21

ULT 4  
68484

ADB68484 standard; peptide; 21 AA.

ADB68484;

04-DEC-2003 (first entry)

Peptide sequence of Pep-2 peptide transfection agent.

hydroxyproline nucleic acid; HypNA; PNA; peptide nucleic acid;  
gene expression; respiration; secretion; signalling;  
ion-channel activity; cell motility; developmental phenotype;  
tumour regression; peptide transfection agent; Pep-2.

Unidentified.

WO2003068798-A2.

21-AUG-2003.

07-FEB-2003; 2003WO-US003904.

09-FEB-2002; 2002US-00072975.

(ACTI-) ACTIVE MOTIF.

Efimov V, Fernandez J, Archdeacon D, Archdeacon J, Choob M;

WPI; 2003-689653/65.

Method of inhibiting expression of genes or RNA transcripts, useful for therapy and determining effects of genes, by administering oligomers containing hydroxyproline nucleic acid.

Claim 18; Page 184; 240pp; English.

The invention relates to a novel method of inhibiting the expression of one or more genes or RNA transcripts by administering at least one oligonucleotide analogue that includes at least one hydroxyproline nucleic acid (HypNA) monomer to a cell or organism or their extracts. The oligonucleotides of the invention may be used to monitor properties including gene expression, respiration, secretion, signalling, ion-channel activity, cell motility, developmental phenotype and tumour regression. Furthermore, they may be utilised to determine the effects of particular genes, as antisense or homologous recombination constructs e.g. for creating animal models of disease and finally, for increasing the activity of some enzymes, such as polymerases. The current sequence is that of the Pep-2 peptide transfection agent of the invention.

SQ Sequence 21 AA;

Query Match 100.0%; Score 135; DB 7; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.3e-10;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KETWETWETWSQPKKRV 21  
|||||  
Db 1 KETWETWETWSQPKKRV 21

RESULT 5

ADC22460  
ID ADC22460 standard; peptide; 21 AA.

XX AC ADC22460;

XX DT 18-DEC-2003 (first entry)

XX Protein-derived transport peptide SEQ ID NO:309.

XX DE recombinant fusion protein; fusion protein; binding; detection;  
KW localisation domain; binding domain;  
KW subcellular compartment localisation.

XX OS Synthetic.

XX PN WO2003012068-A2.

XX PD 13-FEB-2003.

XX PF 01-AUG-2002; 2002WO-US024572.

XX PR 01-AUG-2001; 2001US-0309395P.

XX PR 13-DEC-2001; 2001US-0341589P.

XX PA (CELL-) CELLONICS INC.

XX PI Bright G, Premkumar DR, Chen Y;

XX WPI; 2003-248174/24.

XX New recombinant fusion protein comprising detection and first localisation domains and a binding domain for the molecule of interest, useful for detecting binding of a molecule of interest.

XX Disclosure; SEQ ID NO 309; 101pp; English.

XX The present invention describes a recombinant fusion protein (I) for detecting binding of a molecule of interest. (I) comprises: (a) a detection domain; (b) a first localisation domain; and (c) a binding domain for the molecule of interest. The detection domain, the first localisation domain and the binding domain for the molecule of interest constituting the recombinant fusion protein for detecting binding of a molecule of interest are operably linked. The binding domain for the molecule of interest is separated from the first localisation domain by 0-20 amino acid residues. The first localisation domain and the binding domain for the molecule of interest both do not occur in a single non-recombinant protein with the same spacing as in the recombinant fusion protein for detecting binding of a molecule of interest. Also described: (1) a recombinant nucleic acid encoding the recombinant fusion protein; (2) a recombinant expression vector comprising the nucleic acid control sequences operably linked to the recombinant nucleic acid molecule; (3) a genetically engineered host cell transfected with the recombinant expression vector; (4) a kit for detecting binding of the molecule of interest; and (5) a method for identifying compounds that alter the binding of the molecule of interest. The recombinant fusion protein is useful for detecting binding of a molecule of interest. The recombinant fusion protein eliminates the need to construct two or more chimeric proteins and enables the monitoring of biochemical events in live, intact or fixed cells. The present sequence is used in the exemplification of the present invention.

XX





1 KETWETWTWTEASQPKRK 19

SULT 10  
B77694

ABB77694 standard; peptide; 20 AA.

ABB77694;

01-JUL-2002 (first entry)

Peptide transfection agent Pep-2.10-Cya.

Intracellular delivery; transfection agent; cancer; infectious disease; peptide vector.

Synthetic.

WO200210201-A2.

07-FEB-2002.

26-JUL-2001; 2001WO-US023406.

31-JUL-2000; 2000US-0221932P.

(ACTI-) ACTIVE MOTIF.

(CNRS ) CENT NAT RECH SCI.

Divida G, Morris M, Mery J, Heitz F, Fernandez J, Archdeacon J; Horndorp K;

WPI; 2002-329441/36.

Transfection agent that comprises a peptide comprising hydrophobic and hydrophilic domain and having amino acid residues of specified length is useful for a non-covalent association with and transport of a heterologous compound into a cell.

Example 4; Page 77; 156pp; English.

The invention relates to a transfection agent comprises a peptide of about 16 - 30 amino acids in length. Peptides of the invention comprise a hydrophobic domain, a hydrophilic domain, optionally a spacer sequence between the domains and a functional group conjugated to at least one terminal of the peptide. Peptides of the invention are useful for a non-covalent association with and transport of a heterologous compound into a cell. They are also useful for promoting the cellular internalisation of at least one member e.g. peptide, proteins, antibodies, their derivatives and/or conjugates. They may form part of a pharmaceutical composition to deliver the compound selected from a diagnostic or therapeutic compound, to treat at least one condition such as cancer or an infectious disease, or which targets a cancerous cell or pathogen-infected cell and to deliver a peptide or inhibitor that disrupts the activity of the enzyme. The agent of the invention has a transfection efficiency of at least 5% for at least two of the members of the group of the compounds. The agent has a good delivery efficiency for a broad spectrum of compounds and cell types, has a low toxicity, are easy to handle and easy to formulate in conjunction with the many different compound types that it can deliver. The peptides are serum sensitive, thus they bode particularly well for systemic and/or localised in patients. The current sequence represents the peptide transfection agent Pep-2.10-Cya

Sequence 20 AA;

Query Match 78.1%; Score 105.5; DB 5; Length 20;  
Best Local Similarity 85.7%; Pred. No. 1.1e-06;  
Matches 18; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

1 KETWETWTWTEASQPKRK 21

1 KETWETWTWTEWSQ-KRK 20

RESULT 11

ABB77675

ID ABB77675 standard; peptide; 19 AA.

XX

AC ABB77675;

XX

DT 01-JUL-2002 (first entry)

XX

Peptide transfection agent Pep-2.4.

XX

Intracellular delivery; transfection agent; cancer; infectious disease; peptide vector.

XX

OS Synthetic.

XX

FN WO200210201-A2.

XX

PD 07-FEB-2002.

XX

26-JUL-2001; 2001WO-US023406.

XX

31-JUL-2000; 2000US-0221932P.

XX

(ACTI-) ACTIVE MOTIF.

PA (CNRS ) CENT NAT RECH SCI.

XX

Divida G, Morris M, Mery J, Heitz F, Fernandez J, Archdeacon J;

XX

Horndorp K;

XX

WPI; 2002-329441/36.

XX

Transfection agent that comprises a peptide comprising hydrophobic and

PT hydrophilic domain and having amino acid residues of specified length is

PT useful for a non-covalent association with and transport of a

PT heterologous compound into a cell.

XX

Claim 12; Page 17; 156pp; English.

XX

The invention relates to a transfection agent comprises a peptide of about 16 - 30 amino acids in length. Peptides of the invention comprise a hydrophobic domain, a hydrophilic domain, optionally a spacer sequence between the domains and a functional group conjugated to at least one terminal of the peptide. Peptides of the invention are useful for a non-covalent association with and transport of a heterologous compound into a cell. They are also useful for promoting the cellular internalisation of at least one member e.g. peptide, proteins, antibodies, their derivatives and/or conjugates. They may form part of a pharmaceutical composition to deliver the compound selected from a diagnostic or therapeutic compound, to treat at least one condition such as cancer or an infectious disease, or which targets a cancerous cell or pathogen-infected cell and to deliver a peptide or inhibitor that disrupts the activity of the enzyme. The agent of the invention has a transfection efficiency of at least 5% for at least two of the members of the group of the compounds. The agent has a good delivery efficiency for a broad spectrum of compounds and cell types, has a low toxicity, are easy to handle and easy to formulate in conjunction with the many different compound types that it can deliver. The peptides are serum sensitive, thus they bode particularly well for systemic and/or localised in patients. The current sequence represents the peptide transfection agent Pep-2.4

Sequence 19 AA;

Query Match 75.6%; Score 102; DB 5; Length 19;  
Best Local Similarity 85.7%; Pred. No. 2.9e-06;  
Matches 18; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 KETWETWTWTEASQPKRK 21

DB 1 KETWETW--TWSQPKRK 19

RESULT 12

10106

ABB10106 standard; peptide; 19 AA.

ABB10106;

01-JUL-2002 (first entry)

Peptide transfection agent of the invention.

Intracellular delivery; transfection agent; cancer; infectious disease; peptide vector.

Synthetic.

WO200210201-A2.

07-FEB-2002.

26-JUL-2001; 2001WO-US023406.

31-JUL-2000; 2000US-0221932P.

(ACTI-) ACTIVE MOTIF.  
(CNRS ) CENT NAT RECH SCI.Divida G, Morris M, Mery J, Heitz F, Fernandez J, Archdeacon J;  
Horndorp K;

WPI; 2002-329441/36.

Transfection agent that comprises a peptide comprising hydrophobic and hydrophilic domain and having amino acid residues of specified length is useful for a non-covalent association with and transport of a heterologous compound into a cell.

Disclosure; Page 154; 156pp; English.

The invention relates to a transfection agent comprises a peptide of about 16 - 30 amino acids in length. Peptides of the invention comprise a hydrophobic domain, a hydrophilic domain, optionally a spacer sequence between the domains and a functional group conjugated to at least one terminal of the peptide. Peptides of the invention are useful for a non-covalent association with and transport of a heterologous compound into a cell. They are also useful for promoting the cellular internalisation of at least one member e.g. peptide, proteins, antibodies, their derivatives and/or conjugates. They may form part of a pharmaceutical composition to deliver the compound selected from a diagnostic or therapeutic compound, to treat at least one condition such as cancer or an infectious disease, or which targets a cancerous cell or pathogen-infected cell and to deliver a peptide or inhibitor that disrupts the activity of the enzyme. The agent of the invention has a transfection efficiency of at least 5% for at least two of the members of the group of the compounds. The agent has a good delivery efficiency for a broad spectrum of compounds and cell types, has a low toxicity, are easy to handle and easy to formulate in conjunction with the many different compound types that it can deliver. The peptides are serum sensitive, thus they bode particularly well for systemic and/or localised in patients. The current sequence represents the peptide transfection agent of the invention. Note: this sequence is present in the sequence listing only, and is not referred to at any point in the specification

Sequence 19 AA;

Query Match	73.3%;	Score 99;	DB 5;	Length 19;
Best Local Similarity	72.2%;	3;	Mismatches 2;	Indels 0;
Matches 13;	Conservative	0;	Mismatches 0;	Gaps 0;

2 ETWETWTWTSQPKKR 19  
:|||||:|||||:  
1 KTWETWTWTSQPKKR 18

ULT 13

ABB77689

ID ABB77689 standard; peptide; 12 AA.

XX ABB77689;

XX 01-JUL-2002 (first entry)

XX Hydrophobic Trp-rich motif for design of the Pep-2 vector.

XX Intracellular delivery; transfection agent; cancer; infectious disease; peptide vector.

XX Synthetic.

XX WO200210201-A2.

XX 07-FEB-2002.

XX 26-JUL-2001; 2001WO-US023406.

XX 31-JUL-2000; 2000US-0221932P.

(ACTI-) ACTIVE MOTIF.  
(CNRS ) CENT NAT RECH SCI.Divida G, Morris M, Mery J, Heitz F, Fernandez J, Archdeacon J;  
Horndorp K;

XX WPI; 2002-329441/36.

Transfection agent that comprises a peptide comprising hydrophobic and hydrophilic domain and having amino acid residues of specified length is useful for a non-covalent association with and transport of a heterologous compound into a cell.

Example 3; Page 71; 156pp; English.

The invention relates to a transfection agent comprises a peptide of about 16 - 30 amino acids in length. Peptides of the invention comprise a hydrophobic domain, a hydrophilic domain, optionally a spacer sequence between the domains and a functional group conjugated to at least one terminal of the peptide. Peptides of the invention are useful for a non-covalent association with and transport of a heterologous compound into a cell. They are also useful for promoting the cellular internalisation of at least one member e.g. peptide, proteins, antibodies, their derivatives and/or conjugates. They may form part of a pharmaceutical composition to deliver the compound selected from a diagnostic or therapeutic compound, to treat at least one condition such as cancer or an infectious disease, or which targets a cancerous cell or pathogen-infected cell and to deliver a peptide or inhibitor that disrupts the activity of the enzyme. The agent of the invention has a transfection efficiency of at least 5% for at least two of the members of the group of the compounds. The agent has a good delivery efficiency for a broad spectrum of compounds and cell types, has a low toxicity, are easy to handle and easy to formulate in conjunction with the many different compound types that it can deliver. The peptides are serum sensitive, thus they bode particularly well for systemic and/or localised in patients. The current sequence represents a hydrophobic Trp-rich motif for design of the Pep-2 vector (see ABB77671). This motif is required for efficient targeting to the cell membrane and for forming hydrophobic interactions with proteins

Sequence 12 AA;

Query Match	66.7%;	Score 90;	DB 5;	Length 12;
Best Local Similarity	100.0%;	Pred. No. 5.8e-05;		
Matches 12;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps				

QY 1 KETWETWTWTEW 12  
:|||||:  
DB 1 KETWETWTWTEW 12

RESULT 14

P56174  
 ABP56174 standard; peptide; 12 AA.  
 ABP56174;  
 28-MAR-2003 (first entry)  
 Cell-targeting (TARG) peptide pep-1.  
 Mitochondrial membrane permeabilisation; mitochondrion; PTPC;  
 permeability transition pore complex; viricide; neuroprotective;  
 vasotropic; cytostatic; infection; cell death regulation; apoptosis;  
 mitochondrial permeability transition pore complex modulator; cancer;  
 apoptogenic; ischaemia; neurodegenerative disease; fulminant hepatitis.  
 Synthetic.  
 WO200261105-A2.  
 08-AUG-2002.  
 01-FEB-2002; 2002WO-EP001633.  
 02-FEB-2001; 2001US-0265594P.  
 (INSP ) INST PASTEUR.  
 (CNRS ) CENT NAT RECH SCI.  
 Edelman L, Jacotot E, Briand J;  
 WPI; 2002-619260/66.  
 New chimeric bifunctional molecules that target specific cells and  
 regulate the apoptosis function of the permeability transition pore  
 complex of the mitochondria, useful for treating or preventing e.g.  
 cancer or ischemia.  
 Claim 6; Page 11; 76pp; English.  
 The present invention describes a chimeric bifunctional molecule (I)  
 comprising at least a first functional molecule covalently linked to a  
 second functional molecule, which is able to modulate the activity of the  
 permeability transition pore complex (PTPC) of the mitochondria. (I) has  
 the function of specifically targeting and entering a tissue cell  
 population. The second functional molecule has the function of  
 specifically targeting, and inducing or preventing the death of the cells  
 by apoptosis by regulating the opening or the closing of the PTPC of the  
 mitochondria or its fragment. (I) has viricide, neuroprotective,  
 vasotropic and cytostatic activities, and can be used as a mitochondrial  
 permeability transition pore complex (PTPC) modulator. (I) is useful for  
 treating or preventing a pathological infection or disease. (I) is also  
 useful for regulating cell death regulatory molecules, specifically the  
 apoptogenic function of the PTPC, for treating e.g. cancer, ischaemia,  
 neurodegenerative diseases, fulminant hepatitis or viral infections. The  
 present sequence represents a cell-targeting peptide which is given in  
 the exemplification of the present invention  
 Sequence 12 AA;  
 Query Match 66.7%; Score 90; DB 5; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-05;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 KETWETWTEW 12  
 1 KETWETWTEW 12  
 RESULT 15  
 BB77678  
 ABB77678 standard; peptide; 21 AA.  
 ABB77678;  
 01-JUL-2002 (first entry)  
 Generic peptide transfection agent#2.  
 Intracellular delivery; transfection agent; cancer; infectious disease;  
 peptide vector.  
 Synthetic.  
 Key Location/Qualifiers  
 Misc-difference 2 /label= Xaa  
 /note= "Xaa may be any or no amino acid residue"  
 Misc-difference 3 /label= Xaa  
 /note= "Xaa may be any or no amino acid residue"  
 Misc-difference 10 /label= Xaa  
 /note= "Xaa may be any or no amino acid residue"  
 Misc-difference 11 /label= Xaa  
 /note= "Xaa may be any or no amino acid residue"  
 Misc-difference 12 /label= Xaa  
 /note= "Xaa may be any or no amino acid residue"  
 Misc-difference 18 /label= Xaa  
 /note= "Xaa may be any or no amino acid residue"  
 Misc-difference 21 /label= Xaa  
 /note= "Xaa may be any or no amino acid residue"  
 WO200210201-A2.  
 07-FEB-2002.  
 26-JUL-2001; 2001WO-US023406.  
 31-JUL-2000; 2000US-0221932P.  
 (ACTI-) ACTIVE MOTIF.  
 (CNRS ) CENT NAT RECH SCI.  
 Divida G, Morris M, Mery J, Heitz F, Fernandez J, Archdeacon J;  
 Horndorp K;  
 WPI; 2002-329441/36.  
 Transfection agent that comprises a peptide comprising hydrophobic and  
 hydrophilic domain and having amino acid residues of specified length is  
 useful for a non-covalent association with and transport of a  
 heterologous compound into a cell.  
 Claim 14; Page 18; 156pp; English.  
 The invention relates to a transfection agent comprises a peptide of  
 about 16 - 30 amino acids in length. Peptides of the invention comprise a  
 hydrophobic domain, a hydrophilic domain, optionally a spacer sequence  
 between the domains and a functional group conjugated to at least one  
 terminal of the peptide. Peptides of the invention are useful for a non-  
 covalent association with and transport of a heterologous compound into a  
 cell. They are also useful for promoting the cellular internalisation of  
 at least one member e.g. peptide, proteins, antibodies, their derivatives  
 and/or conjugates. They may form part of a pharmaceutical composition to  
 deliver the compound selected from a diagnostic or therapeutic compound,  
 to treat at least one condition such as cancer or an infectious disease,  
 or which targets a cancerous cell or pathogen-infected cell and to  
 deliver a peptide or inhibitor that disrupts the activity of the enzyme.  
 The agent of the invention has a transfection efficiency of at least 5%  
 for at least two of the members of the group of the compounds. The agent  
 has a good delivery efficiency for a broad spectrum of compounds and cell  
 types, has a low toxicity, are easy to handle and easy to formulate in



conjunction with the many different compound types that it can deliver.  
The peptides are serum sensitive, thus they bode particularly well for  
systemic and/or localised in patients. The current sequence represents a  
generic peptide transfection agent

Sequence 21 AA;

Very Match 66.7%; Score 90; DB 5; Length 21;  
est Local Similarity 70.0%; Pred. No. 0.0001;  
atches 14; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
1 KETWETWETWESQPKKKRK 20  
1 KXXWETWXXXXSQPKKKRK 20

rch completed: February 25, 2004, 14:46:57  
time : 57 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on: February 25, 2004, 14:43:43 ; Search time 21 Seconds  
(without alignments)  
96.192 Million cell updates/sec

le: US-09-915-914B-7

fect score: 135  
uence: 1 KETWETWETWWSQPKKRV 21

ring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

ched: 283366 seqs, 96191526 residues

al number of hits satisfying chosen parameters: 283366

imum DB seq length: 0

imum DB seq length: 2000000000

t-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase : PIR\_78:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

ult No.	Score	Query Match	Length	DB ID	Description
1	61.5	45.6	559	2 B47175	reverse transcript
2	61.5	45.6	1003	1 GNVWLV	HIV-1 retropepsin
3	61.5	45.6	1003	1 B44001	HIV-1 retropepsin
4	61.5	45.6	1003	2 T03440	pol polyprotein -
5	61.5	45.6	1012	1 GNVWLV	HIV-1 retropepsin
6	61.5	45.6	1015	1 GNVWH3	HIV-1 retropepsin
7	56.5	41.9	902	2 T01668	pol polyprotein -
8	56.5	41.9	1002	2 S54378	pol polyprotein -
9	56.5	41.9	1451	2 B86286	F9L1.15 protein -
10	56.5	41.9	1459	2 H96622	probable ABC trans
11	56	41.5	891	1 TVVPAS	large T antigen -
12	56	41.5	895	1 TVVPTB	large T antigen -
13	55.5	41.1	1002	1 GNVJND	HIV-1 retropepsin
14	54.5	40.4	688	1 TVVPTJ	large T antigen -
15	53.5	39.6	708	1 TVVPT4	large T antigen -
16	53.5	39.6	894	2 T27007	hypothetical prote
17	53	39.3	455	2 T48973	cytochrome P450-11
18	53	39.3	689	2 F83902	beta-galactosidase
19	52	38.5	413	2 C91229	hypothetical prote
20	52	38.5	413	2 B86076	hypothetical prote
21	52	38.5	418	2 S40824	hypothetical 48K p
22	52	38.5	710	2 T25734	hypothetical prote
23	50.5	37.4	111	2 AH2530	hypothetical prote
24	50.5	37.4	559	2 A47175	reverse transcript
25	50.5	37.4	812	2 S33980	pol polyprotein -
26	50.5	37.4	1003	1 GNVWA2	HIV-1 retropepsin
27	50	37.0	368	2 T15492	hypothetical prote
28	50	37.0	581	2 T38501	hypothetical prote
29	50	37.0	686	2 AH0104	beta-galactosidase

conserved hypothet  
hypothetical prote  
hypothetical prote  
ecdysone-inducible  
chitinase VCRA0027  
ecdysone-inducible  
poly(3-hydroxyalka  
glycogen phosphory  
hypothetical prote  
protein CTRP - mal  
hypothetical prote  
probable membrane  
hypothetical prote  
hypothetical prote  
hypothetical prote  
interleukin-9 rece

ALIGNMENTS

RESULT 1

B47175 reverse transcriptase, AZT-sensitive variant - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 13-Sep-1996  
C:Accession: B47175  
R:Mohtri, H.; Singh, M.K.; Ching, W.T.; Ho, D.D.  
Proc. Natl. Acad. Sci. U.S.A. 90, 25-29, 1993  
A:Title: Quantitation of zidovudine-resistant human immunodeficiency virus type 1 in the  
A:Reference number: A47175; MUID:93126353; PMID:7678340  
A:Accession: B47175  
A>Status: Preliminary; nucleic acid sequence not shown; not compared with conceptual trar  
A:Molecule type: DNA  
A:Residues: 1-559 <MOH>  
A:Note: Sequence extracted from NCBI backbone (NCBIP:122099)  
C:Superfamily: pol polyprotein

Query Match 45.6%; Score 61.5; DB 2; Length 559;  
Best Local Similarity 78.6%; Pred. No. 2.8;  
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 KETWETWETWWSQ 14

Db 395 KETWETWETWYQ 407  
|||||:|

RESULT 2

GNVWL  
HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 1 (isolate LAV-1a)  
N:Contains: endonuclease (EC 3.1.-.-); retropepsin (EC 3.4.23.16); RNA-directed DNA polym  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 03-Jun-2002  
C:Accession: A03966  
R:Wain-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.  
Cell 40, 9-17, 1985  
A:Title: Nucleotide sequence of the AIDS virus, LAV.  
A:Reference number: A30866; MUID:85099333; PMID:2981635  
A:Accession: A03966  
A:Molecule type: DNA  
A:Residues: 1-1003 <WAL>  
C:Comment: Specific enzymatic cleavages may yield mature proteins including protease, rev  
C:Genetics:  
A:Gene: pol  
C:Superfamily: pol polyprotein  
C:Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucleot  
F:57-155/Product: retropepsin #status predicted <RTP>  
F:81/Active site: Asp (shared with dimeric partner) #status experimental

Query Match 45.6%; Score 61.5; DB 1; Length 1003;  
Best Local Similarity 78.6%; Pred. No. 4.9;  
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

C/Accession: A03967  
 R/Muesing, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.  
 Nature 313, 450-458, 1985  
 A/Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retroviral  
 A/Reference number: A93355; MUID:85111157; PMID:2982104  
 A/Accession: A03967  
 A/Molecule type: DNA  
 A/Residues: 1-1012 <MUE>  
 C/Comment: Specific enzymatic cleavages may yield mature proteins including protease, rev  
 C/Genetics:  
 A/Gene: pol  
 C/Superfamily: pol polyprotein  
 C/Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucleot  
 F/60-164/Product: retropepsin #status predicted <RTP>  
 F/90/Active site: Asp (shared with dimeric partner) #status experimental

Query Match 45.6%; Score 61.5; DB 1; Length 1012;  
 Best Local Similarity 78.6%; Pred. No. 4.9;  
 Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 KETWETWTEWSQ 14  
 DB 559 KET-WETWTEYMQ 571

RESULT 6  
 GNVWH3  
 HIV-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 1 (isolate HTLV-III,  
 N/Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA polym  
 C/Species: human immunodeficiency virus type 1, HIV-1  
 A/Note: host Homo sapiens (man)  
 C/Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 03-Jun-2002  
 C/Accession: A03965  
 R/Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Dorar  
 nberger, J.A.; Papas, T.S.; Chrayeb, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.  
 Nature 313, 277-284, 1985  
 A/Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.  
 A/Reference number: A93353; MUID:85111123; PMID:2578615  
 A/Accession: A03965  
 A/Molecule type: DNA  
 A/Residues: 1-1015 <RAT>  
 A/Cross-references: GB:M15654; GB:K02008; GB:K02009; GB:K02010; NID:G326383; PIDN:AAA441;  
 C/Comment: Specific enzymatic cleavages may yield mature proteins including protease, rev  
 C/Genetics:  
 A/Gene: pol  
 C/Superfamily: pol polyprotein  
 C/Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucleot  
 F/69-167/Product: retropepsin #status predicted <RTP>  
 F/93/Active site: Asp (shared with dimeric partner) #status experimental

Query Match 45.6%; Score 61.5; DB 1; Length 1015;  
 Best Local Similarity 78.6%; Pred. No. 4.9;  
 Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 KETWETWTEWSQ 14  
 DB 562 KET-WETWTEYMQ 574

RESULT 7  
 T01668  
 pol polyprotein - human immunodeficiency virus type 1  
 C/Species: human immunodeficiency virus type 1, HIV-1  
 C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
 C/Accession: T01668  
 R/Alizon, M.; Wain-Hobson, S.; Gluckman, J.C.; Sonigo, P.  
 Cell 46, 63-74, 1986  
 A/Title: Genetic variability of the AIDS virus: Nucleotide sequence analysis of two isol  
 A/Reference number: Z14389; MUID:86245056; PMID:2424612  
 A/Accession: T01668  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1-902 <ALI>

1 KETWETWTEWSQ 14  
 550 KET-WETWTEYMQ 562

SULT 3  
 4001  
 V-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 1 (strain YU-2)  
 Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly  
 Species: human immunodeficiency virus type 1, HIV-1  
 Note: host Homo sapiens (man)  
 Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 03-Jun-2002  
 Accession: B44001  
 Li, Y.; Hui, H.; Burgess, C.J.; Price, R.W.; Sharp, P.M.; Hahn, B.H.; Shaw, G.M.  
 Virol. 56, 6587-6600, 1992  
 Title: Complete nucleotide sequence, genome organization, and biological properties of  
 Reference number: A44001; MUID:93021387; PMID:1404605  
 Accession: B44001  
 Molecule type: DNA  
 Residues: 1-1003 <LIY>  
 Cross-references: GB:M93258  
 Comment: This protein is synthesized as a gag-pol polyprotein.  
 Comment: Specific enzymatic cleavages may yield mature proteins including proteinase,  
 Genetics:  
 Gene: pol  
 Superfamily: pol polyprotein  
 Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucleot  
 S/155/Product: retropepsin #status predicted <RTP>  
 S/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 45.6%; Score 61.5; DB 1; Length 1003;  
 Best Local Similarity 78.6%; Pred. No. 4.9;  
 Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

1 KETWETWTEWSQ 14  
 550 KET-WETWTEYMQ 562

SULT 4  
 9440  
 1 polyprotein - human immunodeficiency virus type 1 (strain JRFL) (fragment)  
 Species: human immunodeficiency virus type 1, HIV-1  
 Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 11-Jan-2000  
 Accession: T09440  
 Pang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazie,  
 bitted to the EMBL Data Library, July 1996  
 Reference number: Z16673  
 Accession: T09440  
 Status: preliminary; translated from GB/EMBL/DBJ  
 Molecule type: DNA  
 Residues: 1-1003 <PAN>  
 Cross-references: EMBL:U63632; NID:G1465777; PID:G1465779  
 Genetics:  
 Gene: pol  
 Superfamily: pol polyprotein

Query Match 45.6%; Score 61.5; DB 2; Length 1003;  
 Best Local Similarity 78.6%; Pred. No. 4.9;  
 Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

1 KETWETWTEWSQ 14  
 550 KET-WETWTEYMQ 562

SULT 5  
 VVYL  
 V-1 retropepsin (EC 3.4.23.16) - human immunodeficiency virus type 1 (isolate LV)  
 Contains: endonuclease (EC 3.1.1.-); retropepsin (EC 3.4.23.16); RNA-directed DNA poly  
 Species: human immunodeficiency virus type 1, HIV-1  
 Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 03-Jun-2002

```

Cross-references: EMBL:K03456; NID:G60228; PIDN:CAA28012.1; PID:G60230
Superfamily: pol polyprotein

Query Match      41.9%; Score 56.5; DB 2; Length 902;
Best Local Similarity 71.4%; Pred. No. 18;
Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

1 KETWETWTEWSQ 14
|||||
449 KET-WEAWTEYQ 461

RESULT 10
H96622
Probable ABC transporter F23H11.19 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: H96622
R:theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: H96622
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1469 <STO>
A:Cross-references: GB:AE005173; NID:G5080820; PIDN:AA039329.1; GSPDB:GN00141
C:Genetics:
A:Gene: F23H11.19
A:Map position: 1

Query Match      41.9%; Score 56.5; DB 2; Length 1469;
Best Local Similarity 36.4%; Pred. No. 28;
Matches 8; Conservative 2; Mismatches 5; Indels 7; Gaps 1;

QY 1 KETW-----WETWTEWSQ 15
:|||||
DB 1188 QNTWQFKSCLWKQWTVWRSP 1209

RESULT 11
TVVPAS
large T antigen - polyomavirus BK (strain AS)
C:Species: Polyomavirus hominis 1 (polyomavirus BK)
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 16-Jul-1999
C:Accession: A33278
R:avis, J.E.; Walker, D.L.; Gardner, S.D.; Frisque, R.J.
J. Virol. 63, 901-911, 1989
A:Title: Nucleotide sequence of the human polyomavirus AS virus, an antigenic variant of
A:Reference number: A33278; MUID:89095020; PMID:2536111
A:Accession: A33278
A:Molecule type: DNA
A:Residues: 1-891 <TAV>
A:Cross-references: GB:M23132; NID:G332774; PIDN:AAA4878.1; PID:G332775
C:Comment: The DNA sequence was obtained from Genbank, release 61.0.
C:Genetics:
A:Introns: 81/3
C:Superfamily: large T antigen; dnaJ amino-terminal homology
C:Keywords: early protein; glycoprotein
F:12-52/Domain: dnaJ amino-terminal homology #status atypical <DNJ>
F:78,155/Binding site: carboxylate (Asn) (covalent) #status predicted

Query Match      41.5%; Score 56; DB 1; Length 691;
Best Local Similarity 26.7%; Pred. No. 16;
Matches 12; Conservative 4; Mismatches 1; Indels 28; Gaps 2;

QY 5 KETWWT-----EWSQ-----PKKGRV 21
|||||
DB 90 WBSWWSFNEKWDDELFCHEDMFASDEEATADSOHSTPPKKGRV 134

RESULT 12
TVVPB
large T antigen - polyomavirus BK
C:Species: Polyomavirus hominis 1 (polyomavirus BK)
C:Date: 28-Feb-1981 #sequence_revision 28-Feb-1981 #text_change 16-Jul-1999
C:Accession: C01632; C36762; A92981; A03608

```

```
Self, I.; Khoury, G.; Dhar, R.
11 18, 963-977, 1979
Title: The genome of human papovavirus BKV.
Reference number: A03632; MUID:80030082; PMID:229976
Accession: C03632
Molecule type: DNA
Residues: 1-695 <SPI>
Cross-references: GB:V01108; GB:J02038; NID:G60844; PIDN:CAA24300.1; PID:G60849
Experimental source: strain Dunlop
Yang, R.C.A.; Wu, R.
ence 205, 456-462, 1979
Title: BK virus DNA: complete nucleotide sequence of a human tumor virus.
Reference number: A36762; MUID:80058557; PMID:228391
Accession: C36762
Molecule type: DNA
Residues: 1-33,35-259,'N',261-334,'K',336,'V',338-445,'R',447-695 <YAN>
Cross-references: GB:V01109; GB:J02039; NID:G60851; PIDN:CAA24302.1; PID:G60852
Experimental source: strain MM
Yang, R.C.A.; Young, A.; Wu, R.
Virol. 34, 416-430, 1980
Title: BK virus DNA sequence coding for the t and T antigens and evaluation of methods
Reference number: A92981; MUID:80185151; PMID:6246273
Accession: A92981
Molecule type: DNA
Residues: 1-695 <VA2>
Experimental source: strain MM
Superfamily: large T antigen; dnaJ amino-terminal homology
Keywords: early protein; glycoprotein
12-62/Domain: dnaJ amino-terminal homology #status atypical <DNJ>
155/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.5%; Score 56; DB 1; Length 695;
Best Local Similarity 26.7%; Pred. No. 16;
Matches 12; Conservative 4; Mismatches 1; Indels 28; Gaps 2;

5 WETWWT-----EWSQ-----PKKKRKV 21
|||||:|:|
90 WESWSSFNKEDLFCHEMFASDEATDSQHSHTPPPKKKKV 134

SULT 13
LJND
V-1 retropepin (EC 3.4.23.16) - human immunodeficiency virus type 1 (isolate NDK)
Contains: endonuclease (EC 3.1.-.-); retropepin (EC 3.4.23.16); RNA-directed DNA poly
Species: human immunodeficiency virus type 1, HIV-1
Note: host Homo sapiens (man)
Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 03-Jun-2002
Accession: JQ0067
Spire, B.; Sire, J.; Zachar, V.; Rey, F.; Barre-Sinoussi, F.; Galibert, F.; Hampe, A.;
ne 81, 275-284, 1989
Title: Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the human immuno
Reference number: JQ0065; MUID:90034200; PMID:2806917
Accession: JQ0067
Molecule type: DNA
Residues: 1-1002 <SPI>
Cross-references: GB:M27323; NID:G328154; PIDN:AAA44869.1; PID:G328158
Comment: Specific enzymatic cleavages may yield mature proteins including protease, re
Genetics:
Gene: pol
Superfamily: pol polyprotein
Keywords: AIDS; aspartic proteinase; endonuclease; hydrolase; immunodeficiency; nucle
56-154/Product: retropepin #status predicted <RTP>
80/Active site: Asp (shared with dimeric partner) #status predicted

Query Match 41.1%; Score 55.5; DB 1; Length 1002;
Best Local Similarity 71.4%; Pred. No. 26;
Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

1 KETWWTWTWTSQ 14
|||||:|:|
549 KETWETWETWQ 561

RESULT 14
TVVPTJ
Large T antigen - Polyomavirus JC
C/Species: Polyomavirus hominis 2 (polyomavirus JC)
C/Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 16-Jul-1999
C/Accession: A03609
R;Frisque, R.J.; Bream, G.L.; Cannella, M.T.
J. Virol. 51, 458-469, 1984
A;Title: Human polyomavirus JC virus genome.
A;Reference number: A03633; MUID:84268011; PMID:6086957
A;Accession: A03609
A;Molecule type: DNA
A;Residues: 1-688 <PRI>
A;Cross-references: GB:J02226; GB:J02227; EMBL:V01118; NID:G1083524; PIDN:AAA82102.1; PII
C;Superfamily: large T antigen; dnaJ amino-terminal homology
C;Keywords: early protein
F;12-62/Domain: dnaJ amino-terminal homology #status atypical <DNJ>

Query Match 40.4%; Score 54.5; DB 1; Length 688;
Best Local Similarity 25.6%; Pred. No. 25;
Matches 11; Conservative 5; Mismatches 4; Indels 23; Gaps 1;

QY 2 ETWWTWTWTSQ-----PKKKRKV 21
|||||:|:|
91 ESWWTFNFKWEDLFCHEMFASDDENTGSHSTPPPKKKKV 133

RESULT 15
TVVPT4
Large T antigen - rhesus polyomavirus
C/Species: Polyomavirus maccacae 1 (rhesus polyomavirus, SV40)
C/Date: 30-Nov-1980 #sequence_revision 30-Nov-1980 #text_change 12-Dec-1997
C/Accession: B03631; B36763; A03607
R;Reddy, V.B.; Thimmapaya, B.; Dhar, R.; Subramanian, K.N.; Zain, B.S.; Pan, J.; Ghosh,
Science 200, 494-502, 1978
A;Title: The genome of simian virus 40.
A;Reference number: A03631; MUID:78159686; PMID:205947
A;Accession: B03631
A;Molecule type: DNA
A;Residues: 1-708 <RED>
R;Fiers, W.; Contreras, R.; Haegeman, G.; Rogiers, R.; van de Voorde, A.; van Heuverswyn,
Nature 273, 113-120, 1978
A;Title: Complete nucleotide sequence of SV40 DNA.
A;Reference number: A36763; MUID:78156432; PMID:205802
A;Accession: B36763
A;Molecule type: DNA
A;Residues: 1-530,'Y',532-548,'P',550-551,'P',553-708 <PIE>
A;Experimental source: strain 776
C;Comment: The initial 82 residues of the large T and small t antigens from SV40 are code
C;Superfamily: large T antigen; dnaJ amino-terminal homology
C;Keywords: acetylated amino end
F;12-62/Domain: dnaJ amino-terminal homology #status atypical <DNJ>
F;1/Modified site: acetylated amino end (Met) #status experimental

Query Match 39.6%; Score 53.5; DB 1; Length 708;
Best Local Similarity 26.2%; Pred. No. 33;
Matches 11; Conservative 3; Mismatches 3; Indels 25; Gaps 1;

QY 5 WETWWTWTSQ-----PKKKRKV 21
|||||:|:|
91 WEQWNAFNEENLFCSEMPSSDEATDSQHSHTPPPKKKKV 132

Search completed: February 25, 2004, 14:48:48
Job time : 25 secs
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

protein - protein search, using sw model

on: February 25, 2004, 14:41:08 ; Search time 12 Seconds

(without alignments)

91.123 Million cell updates/sec

le: US-09-915-914B-7

fect score: 135

uence: 1 KETWETWTEWSQPKKKRV 21

ring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

ched: 141681 seqs, 52070155 residues

al number of hits satisfying chosen parameters: 141681

imum DB seq length: 0

imum DB seq length: 2000000000

t-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

ult No.	Score	Query Match	Length	DB ID	Description
1	61.5	45.6	1003	1	POL_HV1H2
2	61.5	45.6	1003	1	POL_HV1H2
3	61.5	45.6	1007	1	POL_HV1H2
4	61.5	45.6	1015	1	POL_HV1B1
5	61.5	45.6	1015	1	POL_HV1B5
6	61.5	45.6	1015	1	POL_HV1B8
7	61.5	45.6	1015	1	POL_HV1P
8	60.5	44.8	1006	1	POL_HV1M
9	56.5	41.9	1002	1	POL_HV1E1
10	56.5	41.9	1002	1	POL_HV1A
11	56.5	41.9	1002	1	POL_HV1H
12	56.5	41.9	1002	1	POL_HV1Z
13	56.5	41.9	1003	1	POL_HV1N5
14	56.5	41.9	1003	1	POL_HV1OY
15	56.5	41.9	691	1	TALA_POVEA
16	56.5	41.5	695	1	TALA_POVEB
17	55.5	41.1	1002	1	POL_HV1ND
18	54.5	40.4	688	1	TALA_POVJC
19	53.5	39.6	708	1	TALA_SV40
20	52	38.5	413	1	YHS_ECOLI
21	50.5	37.4	1002	1	POL_HV1U4
22	50.5	37.4	1003	1	POL_HV1A2
23	48.5	35.9	692	1	PHSG_AQUAE
24	48	35.6	41	1	LPW_VIBPA
25	48	35.6	468	1	IL9R_MOUSE
26	48	35.6	586	1	TALA_POVEB
27	48	35.6	1019	1	POL_SIVS4
28	48	35.6	1022	1	POL_SIVSP
29	48	35.6	1054	1	POL_SIVMK
30	48	35.6	1056	1	POL_SIVM1
31	48	35.6	2298	1	YCF2_LOTVA
32	47.5	35.2	315	1	SECF_SYNY3
33	47.5	35.2	355	1	HKM1_LYCES

34 47.5 35.2 398 1 HKL1\_ARATH  
35 47.5 35.2 1027 1 POL\_SIVCZ  
36 47 34.8 373 1 PRXC\_CALFU  
37 47 34.8 633 1 GYP3\_YEAST  
38 46.5 34.4 794 1 YE14\_YEAST  
39 46.5 34.4 1058 1 POL\_HV2D2  
40 46 34.1 78 1 YO09\_BPL2  
41 46 34.1 229 1 UNG\_CHLCV  
42 46 34.1 504 1 MATK\_NEPAL  
43 46 34.1 506 1 MATK\_CALVU  
44 46 34.1 513 1 MATK\_CYRRA  
45 46 34.1 567 1 EIL3\_ARATH

#### ALIGNMENTS

RESULT 1  
POL\_HV1H2  
ID POL\_HV1H2 STANDARD; PRT; 1003 AA.  
AC P04585; Q09777; Q9WJC5;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE POL polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.16);  
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].  
GN POL.  
OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11706;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87299196; PubMed=3040055;  
RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,  
Gallo R.C., Wong-Staal F.;  
RT "Complete nucleotide sequences of functional clones of the AIDS  
virus";  
RL AIDS Res. Hum. Retroviruses 3:57-69(1987).  
RN [2]  
RP REVISIONS.  
RA Ogata N., Alter H.J., Miller R.H., Purcell R.H.;  
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Chappey C.;  
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 57-155.  
RX MEDLINE=93043699; PubMed=9827997;  
RA Kervinen J., Lubkowski J., Zdanov A., Bhatt D., Dunn B.M., Hui K.Y.,  
Powell D.J., Kay J., Wlodawer A., Gustchina A.;  
RT "Toward a universal inhibitor of retroviral proteases: comparative  
analysis of the interactions of LP-130 complexed with proteases from  
HIV-1, FIV, and EIAV";  
RL Protein Sci. 7:2314-2323(1998).  
RN [5]  
RP STRUCTURE BY NMR OF 57-155.  
RX MEDLINE=97022126; PubMed=8868486;  
RA Yamazaki T., Hinch A.P., Wang Y.-X., Nicholson L.K., Torchia D.A.,  
Wingfield P., Stahl S.J., Kaufman J.D., Chang C.-H., Domaille P.J.,  
Lam P.F.S.;  
RT "Three-dimensional solution structure of the HIV-1 protease complexed  
with DMP323, a novel cyclic urea-type inhibitor, determined by  
nuclear magnetic resonance spectroscopy";  
RL Protein Sci. 5:495-506(1996).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 156-595.  
RX MEDLINE=96097398; PubMed=8535785;  
RA Ren J., Esnouf R.M., Hopkins A.L., Ross C., Jones E.Y., Stammers D.K.,  
Stuart D.I.;  
RT "The structure of HIV-1 reverse transcriptase complexed with  
9-chloro-TIBO: lessons for inhibitor design";  
RL Structure 3:915-926(1995).

[7]  
X-RAY CRYSTALLOGRAPHY (2.55 ANGSTROMS) OF 156-595.  
MEDLINE=96208551; PubMed=8648598;  
Hopkins A.L., Ren J., Esnouf R.M., Wilcox B.E., Jones E.Y., Ross C.,  
Miyasaka T., Walker R.T., Tanaka H., Stammers D.K., Stuart D.I.,  
"Complexes of HIV-1 reverse transcriptase with inhibitors of the HEPT  
series reveal conformational changes relevant to the design of potent  
non-nucleoside inhibitors.";  
J. Med. Chem. 39:1589-1600(1996).  
[8]  
X-RAY CRYSTALLOGRAPHY (2.65 ANGSTROMS) OF 157-694.  
MEDLINE=97268693; PubMed=9108091;  
Esnouf R.M., Ren J., Hopkins A.L., Ross C.K., Jones E.Y.,  
Stammers D.K., Stuart D.I.,  
"Unique features in the structure of the complex between HIV-1  
reverse transcriptase and the bis(heteroaryl)piperazine (BHAP)  
U-90152 explain resistance mutations for this nonnucleoside  
inhibitor.";  
Proc. Natl. Acad. Sci. U.S.A. 94:3984-3989(1997).  
[9]  
X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 157-694.  
MEDLINE=98356189; PubMed=9689112;  
Ren J., Esnouf R.M., Hopkins A.L., Jones E.Y., Kirby I., Keeling J.,  
Ross C.K., Larder B.A., Stuart D.I., Stammers D.K.,  
"3'-Azido-3'-deoxythymidine drug resistance mutations in HIV-1  
reverse transcriptase can induce long range conformational changes.";  
Proc. Natl. Acad. Sci. U.S.A. 95:9518-9523(1998).  
-i- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,  
and P1' variable, but often Pro.  
-i- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
phosphononucleoside.  
-i- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate  
+ {DNA}(N).  
-i- PTM: Cleavage sites that yield the mature proteins remain to be  
determined.  
-i- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
KNOWN AS THE RETROPEPSIN FAMILY.  
-----  
THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; K03455; AAB50259.1; ALT\_INIT.  
EMBL; AF033819; AAC82598.2; --  
PDB; 1REV; 14-OCT-96.  
PDB; 1RT1; 21-APR-97.  
PDB; 1RT2; 21-APR-97.  
PDB; 1RT3; 16-FEB-99.  
PDB; 1RTH; 03-APR-96.  
PDB; 1RTU; 03-APR-96.  
PDB; 1BVG; 17-AUG-96.  
PDB; 1BVE; 17-AUG-96.  
PDB; 1KLM; 18-MAR-98.  
PDB; 1A30; 29-APR-98.  
PDB; 1ODI; 16-FEB-99.  
PDB; 1BV7; 14-JAN-00.  
PDB; 1BV9; 12-JAN-00.  
PDB; 1BWA; 12-JAN-00.  
PDB; 1BWB; 12-JAN-00.  
PDB; 1COT; 19-JUL-00.  
PDB; 1COU; 19-JUL-00.  
PDB; 1CIB; 21-JUL-00.  
PDB; 1C1C; 06-SEP-00.  
PDB; 1DMP; 12-NOV-97.  
PDB; 1DTQ; 20-MAR-00.  
PDB; 1DTT; 02-APR-00.  
PDB; 1EP4; 27-SEP-00.  
PDB; 1EX4; 26-JUL-00.  
PDB; 1EXQ; 03-NOV-00.  
PDB; 1FK9; 28-FEB-01.  
PDB; 1FKO; 28-FEB-01.  
PDB; 1FKP; 03-NOV-00.  
PDB; 1HVV; 06-APR-99.  
PDB; 1HVR; 15-MAY-95.  
PDB; 1HWR; 23-MAR-93.  
PDB; 1HXH; 12-MAR-97.  
PDB; 1JXH; 03-OCT-01.  
PDB; 1JLA; 03-OCT-01.  
PDB; 1JLB; 03-OCT-01.  
PDB; 1JLC; 03-OCT-01.  
PDB; 1JLF; 03-OCT-01.  
PDB; 1JLG; 03-OCT-01.  
PDB; 1JLQ; 22-AUG-01.  
PDB; 1LW0; 30-OCT-02.  
PDB; 1LW2; 30-OCT-02.  
PDB; 1LW3; 30-OCT-02.  
PDB; 1LWE; 30-OCT-02.  
PDB; 1LWF; 30-OCT-02.  
PDB; 1MER; 15-APR-98.  
PDB; 1MES; 15-APR-98.  
PDB; 1MET; 15-APR-98.  
PDB; 1MEU; 15-APR-98.  
PDB; 1OIW; 18-FEB-03.  
PDB; 1ODW; 01-APR-97.  
PDB; 1QBR; 15-OCT-97.  
PDB; 1QBS; 15-OCT-97.  
PDB; 1QBT; 15-OCT-97.  
PDB; 1QBU; 15-OCT-97.  
PDB; 1RT4; 29-JUL-99.  
PDB; 1RT5; 29-JUL-99.  
PDB; 1RT6; 29-JUL-99.  
PDB; 1RT7; 29-JUL-99.  
PDB; 1RTD; 12-JAN-00.  
PDB; 1VRT; 03-APR-96.  
PDB; 3PHV; 15-JAN-92.  
PDB; K03455; POLSHX82.  
MEROPS; A02.001; --  
InterPro; IPR001969; Asprotease\_AS.  
InterPro; IPR001037; Integrase\_C.  
InterPro; IPR003308; Integrase\_Zn.  
InterPro; IPR009007; Pept\_A\_acid.  
InterPro; IPR001995; Peptidase\_A2.  
InterPro; IPR002156; RNaseH.  
InterPro; IPR001584; Rve.  
InterPro; IPR000477; RVTse.  
Pfam; PF00552; Integrase; 1.  
Pfam; PF02022; Integrase\_Zn; 1.  
Pfam; PF00075; RNaseH; 1.  
Pfam; PF00655; rve; 1.  
Pfam; PF00077; rvp; 1.  
Pfam; PF00078; rvt; 1.  
PROSITE; PS00141; ASP\_PROTEASE; 1.  
PROSITE; PS00175; ASP\_PROT\_RETROV; 1.  
AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;  
Kw Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.  
PROTEASE.  
CHAIN 57 155  
FT ACT SITE 81 81  
FT STRAND 58 59  
FT STRAND 66 71  
FT TURN 72 73  
FT STRAND 74 80  
FT TURN 82 83  
FT STRAND 88 89  
FT STRAND 99 104  
FT STRAND 109 122  
FT STRAND 126 133  
FT STRAND 140 141  
FT HELIX 143 146  
FT TURN 147 150  
FT STRAND 152 154  
FT STRAND 167 167

DR Pfam: PF00078; rvt; 1.  
 DR PROSITE: PS00141; ASP\_PROTEASE; 1.  
 DR PROSITE: PS00175; ASP\_PROT\_RETROV; 1.  
 DR KW Aids; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;  
 DR Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.  
 DR CHAIN 57 155  
 DR FT ACT SITE 81 BY SIMILARITY.  
 DR SQ SEQUENCE 1003 AA; 113794 MW; D2F7187FE4E8F49 CRC64;

Query Match 45.6%; Score 61.5; DB 1; Length 1003;  
 Best Local Similarity 78.6%; Pred.No.2;  
 Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 KETWETWETWESQ 14  
 ||| |||||:  
 ||| |||||:  
 Db 550 KET-WETWETWYQ 562

RESULT 3  
 ID POL\_HVJUR STANDARD; PRT; 1007 AA.  
 AC P20875;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE POL polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.16);  
 DE Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].  
 GN POL.  
 OS Human immunodeficiency virus type 1 (JRCSP isolate) (HIV-1).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 NCBI\_Taxid=11688;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kovanagi S., Chen I.S.Y.;  
 RL Submitted (DEC-1988) to the HIV data bank.  
 CC -!- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,  
 CC and P1' variable, but often Pro.  
 CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 CC phosphomononucleosides.  
 CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate  
 CC + [DNA] (N).  
 CC -!- PTM: Cleavage sites that yield the mature proteins remain to be  
 CC determined.  
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
 CC KNOWN AS THE RETROPEPSIN FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: M38429; AAB03745.1; -.  
 CC HSP: P03366; 1HMV.  
 CC HIV: M38429; POLSJRCFS.  
 CC MEROPS: A02.001; -.  
 CC InterPro: IPR001969; Aspartase\_AS.  
 CC InterPro: IPR001037; Integrase\_C.  
 CC InterPro: IPR003308; Integrase\_Zn.  
 CC InterPro: IPR009007; Pept\_Acid.  
 CC InterPro: IPR001995; Peptidase\_A2.  
 CC InterPro: IPR002156; RNaseH.  
 CC InterPro: IPR001584; Rve.  
 CC InterPro: IPR000477; RVTse.  
 CC Pfam: PF00552; Integrase; 1.  
 CC Pfam: PF02022; Integrase\_Zn; 1.  
 CC Pfam: PF00075; rnaseH; 1.  
 CC Pfam: PF00665; rve; 1.  
 CC Pfam: PF00077; rvp; 1.  
 CC Pfam: PF00078; rvt; 1.  
 CC PROSITE: PS00141; ASP\_PROTEASE; 1.



PROSITE; PSS0175; ASP\_PROT\_RETROV; 1.  
 AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;  
 Nuclease; Transferase; RNA-directed DNA polymerase.  
 CHAIN 61 159 PROTEASE  
 ACT SITE 85 BY SIMILARITY  
 SEQUENCE 1007 AA; 114081 MW; 51236562EF3411A CRC64;  
 Query Match 45.6%; Score 61.5; DB 1; Length 1007;  
 Best Local Similarity 78.6%; Pred. No. 2;  
 Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
 1 KETWETWETWENSO 14  
 ||| |||||  
 554 KETWETWETWYQ 566  
 SULT 4  
 L\_HV1B1  
 \_POL HV1B1 STANDARD; PRT; 1015 AA.  
 P03366;  
 21-JUL-1986 (Rel. 01, Created)  
 01-NOV-1988 (Rel. 09, Last sequence update)  
 10-OCT-2003 (Rel. 42, Last annotation update)  
 POL polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.16);  
 Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].  
 POL.  
 Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).  
 Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 NCBI\_TaxID=11678;  
 [1]  
 SEQUENCE FROM N.A.  
 MEDLINE=85111123; PubMed=2578615;  
 Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,  
 Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,  
 Baumesister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,  
 Lautenberger J.A., Papas T.S., Chrayeb J., Chang N.T., Gallo R.C.,  
 Wong-Staal F.;  
 "Complete nucleotide sequence of the AIDS virus, HTLV-III.";  
 Nature 313:277-284(1985).  
 [2]  
 3D-STRUCTURE MODELING OF PROTEASE DOMAIN.  
 MEDLINE=89146134; PubMed=2537531;  
 Weber I.T., Miller M., Jaskolski M., Leis J., Skalka A.M.,  
 Wlodawer A.;  
 "Molecular modeling of the HIV-1 protease and its substrate binding  
 site.";  
 Science 243:928-931(1989).  
 [3]  
 X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 69-167.  
 MEDLINE=90044107; PubMed=2682266;  
 Lapatto R., Blundell T., Hemmings A., Overington J., Wilderspin A.,  
 Wood S., Merson J.R., Whittle P.J., Danley D.E., Geoghegan K.F.,  
 Hawrylik S.J., Lee S.E., Scheid K.G., Hobart P.M.;  
 "X-ray analysis of HIV-1 protease at 2.7-A resolution confirms  
 structural homology among retroviral enzymes.";  
 Nature 342:299-302(1989).  
 [4]  
 X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 69-167.  
 MEDLINE=90341771; PubMed=2200122;  
 Erickson J., Neidhart D.J., Vandrie J., Kempf D.J., Wang X.C.,  
 Norbeck D.W., Plattner J.J., Rittenhouse J.W., Turon M., Wideburg N.,  
 Kohlbrener W.E., Simmer R., Helfrich R., Paul D.A., Knigge M.;  
 "Design, activity, and 2.8 A crystal structure of a C2 symmetric  
 inhibitor complexed to HIV-1 protease.";  
 Science 249:527-533(1990).  
 [5]  
 X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 594-729.  
 MEDLINE=91188281; PubMed=1707186;  
 Davies J.F. II, Hostomska Z., Hostomsky Z., Jordan S.R.,  
 Matthews D.A.;  
 "Crystal structure of the ribonuclease H domain of HIV-1 reverse  
 transcriptase.";  
 Science 252:88-95(1991).

RN X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 168-723.  
 RX MEDLINE=93317673; PubMed=7687065;  
 RA Jacobo-Molina A., Ding J., Nanni R.G., Clark A.D. Jr., Lu X.,  
 Tantillo C., Williams R.L., Kamer G., Ferris A.L., Clark P., Hizi A.,  
 Hughes S.H., Arnold E.;  
 RT "Crystal structure of human immunodeficiency virus type 1 reverse  
 transcriptase complexed with double-stranded DNA at 3.0-A resolution  
 shows bent DNA.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6320-6324(1993).  
 RN [7]  
 RN X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS) OF 168-723.  
 RX MEDLINE=92311654; PubMed=1377403;  
 RA Kohlstaedt L.A., Wang J., Friedman J.M., Rice P.A., Steitz T.A.;  
 RT "Crystal structure at 3.5-A resolution of HIV-1 reverse transcriptase  
 complexed with an inhibitor.";  
 RL Science 256:1783-1790(1992).  
 RN [8]  
 RN X-RAY CRYSTALLOGRAPHY (3.2 ANGSTROMS) OF 168-727.  
 RX MEDLINE=95166801; PubMed=7532306;  
 RA Rodgers D.W., Gamblin S.J., Harris B.A., Ray S., Culp J.S.,  
 Hellwig B., Woolf D.J., Debouck C., Harrison S.C.;  
 RT "The structure of unliganded reverse transcriptase from the human  
 immunodeficiency virus type 1.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:1222-1226(1995).  
 RN [9]  
 RN X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 168-725.  
 RX MEDLINE=95338599; PubMed=7542140;  
 RA Ding J., Das K., Tantillo C., Zhang W., Clark A.D. Jr., Jessen S.,  
 Lu X., Hsiou Y., Jacobo-Molina A., Andries K., Et A.L.;  
 RT "Structure of HIV-1 reverse transcriptase in a complex with the non-  
 nucleoside inhibitor alpha-APA R 95845 at 2.8-A resolution.";  
 RL Structure 3:365-379(1995).  
 RN [10]  
 RN X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 168-723.  
 RX MEDLINE=96434330; PubMed=8905568;  
 RA Hsiou Y., Ding J., Das K., Clark A.D. Jr., Hughes S.H., Arnold E.;  
 RT "Structure of unliganded HIV-1 reverse transcriptase at 2.7-A  
 resolution: implications of conformational changes for polymerization  
 and inhibition mechanisms.";  
 RL Structure 4:853-860(1996).  
 RN [11]  
 RN X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 168-723.  
 RX MEDLINE=99033049; PubMed=9813120;  
 RA Hsiou Y., Das K., Ding J., Clark A.D. Jr., Kleim J.P., Rosner M.,  
 Winkler I., Riess G., Hughes S.H., Arnold E.;  
 RT "Structures of Tyr188Leu mutant and wild-type HIV-1 reverse  
 transcriptase complexed with the non-nucleoside inhibitor HBV 097:  
 inhibitor flexibility is a useful design feature for reducing drug  
 resistance.";  
 RL J. Mol. Biol. 284:313-323(1998).  
 CC -|- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,  
 and P1' variable, but often Pro.  
 CC -|- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 phosphonooxester.  
 CC -|- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate  
 + {DNA}(N)  
 CC -|- PTM: Cleavage sites that yield the mature proteins remain to be  
 determined.  
 CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
 KNOWN AS THE RETROPEPSIN FAMILY.  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL: M15654; AAA44198.1; --  
 DR PIR: A03965; GNVH3.  
 DR PDB: 1HVP; 15-APR-92.

PDB; 9HVP; 15-JUL-92.  
PDB; 1HRH; 15-OCT-94.  
PDB; 2HM1; 14-OCT-98.  
PDB; 1HOS; 31-OCT-93.  
PDB; 3HVT; 15-OCT-94.  
PDB; 1HVI; 30-APR-94.  
PDB; 1HVK; 30-APR-94.  
PDB; 1HVL; 30-APR-94.  
PDB; 1HEG; 31-MAY-94.  
PDB; 1HEF; 31-MAY-94.  
PDB; 1HMV; 31-MAR-95.  
PDB; 1HNI; 03-JUN-95.  
PDB; 1HNV; 10-JUL-95.  
PDB; 1HPS; 31-AUG-94.  
PDB; 1HTE; 31-JUL-94.  
PDB; 1HTE; 31-JUL-94.  
PDB; 1HTG; 31-JUL-94.  
PDB; 1BSG; 15-OCT-94.  
PDB; 1DLO; 01-AUG-96.  
PDB; 1GNN; 08-NOV-96.  
PDB; 1GNO; 08-NOV-96.  
PDB; 1AJV; 20-AUG-97.  
PDB; 1AJX; 17-SEP-97.  
PDB; 1MER; 15-APR-98.  
PDB; 1MES; 15-APR-98.  
PDB; 1MET; 15-APR-98.  
PDB; 1MEU; 15-APR-98.  
PDB; 1BQM; 16-FEB-99.  
PDB; 1BQN; 16-FEB-99.  
PDB; 1DAH; 26-JUN-02.  
PDB; 1DAI; 26-JUN-02.  
PDB; 1DAJ; 26-JUN-02.  
PDB; 1DM6; 26-JUL-00.  
PDB; 1EBK; 26-JUL-00.  
PDB; 1EBW; 26-JUN-02.  
PDB; 1EBZ; 26-JUN-02.  
PDB; 1BEC; 26-JUN-02.  
PDB; 1EC1; 26-JUN-02.  
PDB; 1EC2; 26-JUN-02.  
PDB; 1EC3; 26-JUN-02.  
PDB; 1EET; 07-FEB-01.  
PDB; 1EBV; 10-JUL-95.  
PDB; 1HH; 10-JUL-95.  
PDB; 1HPZ; 30-MAY-01.  
PDB; 1HQU; 30-MAY-01.  
PDB; 1HYS; 26-MAR-01.  
PDB; 1IKV; 17-APR-02.  
PDB; 1IKX; 17-APR-02.  
PDB; 1J5O; 14-JUN-02.  
PDB; 1NSY; 28-JAN-03.  
PDB; 1N6Q; 14-JAN-03.  
PDB; 1QMC; 14-DEC-99.  
PDB; 1RTD; 12-JAN-00.  
PDB; 1TVR; 12-MAR-97.  
PDB; 1UWB; 15-MAY-97.  
PDB; 3TLH; 03-FEB-00.  
HIV; M15654; POLSBH102.  
MEROPS; A02.001; -.  
InterPro; IPR001969; Aspartate AS.  
InterPro; IPR001037; Integrase C.  
InterPro; IPR003308; Integrase Zn.  
InterPro; IPR009007; PeptAcid.  
InterPro; IPR001995; Peptidase A2.  
InterPro; IPR002156; RNaseH.  
InterPro; IPR001584; Rve.  
InterPro; IPR000477; RVTse.  
Pfam; PF00552; Integrase; 1.  
Pfam; PF02022; Integrase Zn; 1.  
Pfam; PF00075; rnaaseh; 1.  
Pfam; PF00665; rve; 1.

Query Match 45.6%; Score 61.5; DB 1; Length 1015;  
Best Local Similarity 78.6%; Pred. No. 2;  
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
QY 1 KETWTWETWETWQ 14  
Db 562 KET-WETWTWETWQ 574  
RESULT 5  
POL\_HV1B5  
ID\_POL\_HV1B5 STANDARD; PRT; 1015 AA.  
AC P04587;  
DT 13-AUG-1987 (Rel. 05, Created)  
DT 01-NOV-1988 (Rel. 09, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE POL polypeptide [Contains: Protease (Retropepsin) (EC 3.4.23.16);  
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].  
GN POL.  
OS Human immunodeficiency virus type 1 (BH5 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11682;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85111123; PubMed=2578615;  
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,  
Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,  
Baumister K., Ivanoff L., Pettey S.R., Pearson M.L.,  
Lautenberger J.A., Papas T.S., Gray J., Chang N.T., Gallo R.C.,  
Wong-Staal F.;  
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";  
RL Nature 313:277-284(1985).  
CC -!- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,  
and P1', variable, but often pro.  
CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
phosphomononucleosides.  
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate  
+ {DNA}(N).  
CC -!- PTM: Cleavage sites that yield the mature proteins remain to be  
determined.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
KNOWN AS THE RETROPEPSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; K02012; AAA44653.1; -.  
DR PDB; 1BDL; 12-AUG-98.  
DR PDB; 1BDQ; 12-AUG-98.  
DR PDB; 1BDR; 14-OCT-98.  
DR PDB; 1FEJ; 01-JUN-01.  
DR PDB; 1FEF; 01-JUN-01.  
DR PDB; 1FFF; 01-JUN-01.  
DR PDB; 1FFI; 01-JUN-01.  
DR PDB; 1FG6; 01-JUN-01.  
DR PDB; 1FG8; 01-JUN-01.  
DR PDB; 1FGC; 01-JUN-01.  
DR PDB; 1G2K; 16-NOV-01.  
DR PDB; 1HPV; 31-MAR-95.  
DR PDB; 1HTF; 31-JUL-94.  
DR PDB; 1HVP; 30-APR-94.  
DR PDB; 1HVS; 14-FEB-95.  
DR PDB; 1KI1; 10-JUL-02.  
DR PDB; 1KI1; 10-JUL-02.  
DR PDB; 1K2B; 10-JUL-02.  
DR PDB; 1K2C; 10-JUL-02.  
DR PDB; 1ODX; 01-APR-97.

PDB; 1TCX; 07-DEC-96.  
 PDB; 1WJE; 16-DEC-98.  
 PDB; 1WJF; 16-DEC-98.  
 PDB; 2BPV; 23-FEB-99.  
 PDB; 2BPW; 23-FEB-99.  
 PDB; 2BPX; 23-FEB-99.  
 PDB; 2BPY; 23-FEB-99.  
 PDB; 2BPZ; 23-FEB-99.  
 HIV; K02012; POL\$BHS.

MEROPS; A02.001; -.  
 InterPro; IPR001969; Aspartic protease AS.

InterPro; IPR001037; Integrase C.

InterPro; IPR003308; Integrase Zn.

InterPro; IPR009007; Peptidase A2.

InterPro; IPR001995; Peptidase A2.

InterPro; IPR002156; RNaseH.

InterPro; IPR001584; Rve.

InterPro; IPR000477; RVTse.

Pfam; PF00552; Integrase; 1.

Pfam; PF02022; Integrase Zn; 1.

Pfam; PF00075; RNaseH; 1.

Pfam; PF00665; rve; 1.

Pfam; PF00077; rvp; 1.

Pfam; PF00078; rvt; 1.

PROSITE; PS00141; ASP\_PROTEASE; 1.

PROSITE; PS00175; ASP\_PROT\_RETROV; 1.

AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;

Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.

CHAIN 69 167

ACT\_SITE 93 93

BY SIMILARITY.

SEQUENCE 1015 AA; 115015 MW; F79FCB4A1A92CEE CRC64;

Query Match 45.6%; Score 61.5; DB 1; Length 1015;

Best Local Similarity 78.6%; Pred. No. 2;

Matches 11; Conservative 1; Mismatches 1; Gaps 1;

1 KETWETWTEWSQ 14

|||||

562 KET-WETWTEYQ 574

SULT 6

L\_HVBR

POL\_HVBR

STANDARD;

PRT; 1015 AA.

P03367;

21-JUL-1986 (Rel. 01, Created)

01-NOV-1988 (Rel. 09, Last sequence update)

10-OCT-2003 (Rel. 42, Last annotation update)

POL polyprotein [Contains: Protease (retropepsin) (EC 3.4.23.16);

Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].

POL.

Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).

Viruses; Retroviral viruses; Retroviridae; Lentivirus.

NCBI\_TaxID=11686;

[1]

SEQUENCE FROM N.A.

MEDLINE=8509333; PubMed=2981635;

Wain-Hobson S., Sonigo P., Danos O., Cole S., Alizon M.;

"Nucleotide sequence of the AIDS virus, LAV.,"

Cell 40:9-17(1985).

[2]

REVISTONS TO 23-35.

MEDLINE=86245056; PubMed=2424612;

Alizon M., Wain-Hobson S., Montagnier L., Sonigo P.;

"Genetic variability of the AIDS virus; nucleotide sequence analysis

of two isolates from African patients.,"

Cell 46:63-74(1986).

[3]

X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS).

MEDLINE=92190341; PubMed=1799632;

Spinelli S., Liu Q.Z., Alzari P.M., Harel P.H., Poljak R.J.;

"The three-dimensional structure of the aspartyl protease from the

HIV-1 isolate BRU.,"

Biochimie 73:1391-1396(1991).

-|- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,

and P1' variable, but often Pro.

-|- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-

phosphomonoester.

-|- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate

+{DNA}(N).

-|- PTM: Cleavage sites that yield the mature proteins remain to be

determined.

-|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO

KNOWN AS THE RETROPEPSIN FAMILY.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; K02013; -; NOT ANNOTATED\_CDS.

PDB; 1HHP; 15-OCT-92.

PDB; 1ASG; 15-JUL-98.

PDB; 1A94; 16-FEB-99.

PDB; 1D48; 12-OCT-99.

PDB; 1D4Y; 13-OCT-99.

PDB; 1D4Z; 03-MAY-00.

PDB; 1DIF; 08-MAR-96.

PDB; 1HQS; 31-OCT-93.

PDB; 1HPO; 21-APR-97.

PDB; 1HFS; 15-JAN-95.

PDB; 1HFX; 08-MAR-96.

PDB; 1HSG; 03-APR-96.

PDB; 1HTE; 31-JUL-94.

PDB; 1HVL; 30-APR-94.

PDB; 1UPJ; 14-OCT-96.

PDB; 2URJ; 14-OCT-96.

PDB; 7UFU; 21-APR-97.

DR MEROPS; A02.001; -.

DR InterPro; IPR001969; Aspartic protease AS.

DR InterPro; IPR001037; Integrase C.

DR InterPro; IPR003308; Integrase Zn.

DR InterPro; IPR009007; Peptidase A2.

DR InterPro; IPR001995; Peptidase A2.

DR InterPro; IPR002156; RNaseH.

DR InterPro; IPR001584; Rve.

DR InterPro; IPR000477; RVTse.

Pfam; PF00552; Integrase; 1.

Pfam; PF02022; Integrase Zn; 1.

Pfam; PF00075; RNaseH; 1.

Pfam; PF00665; rve; 1.

Pfam; PF00077; rvp; 1.

Pfam; PF00078; rvt; 1.

DR PROSITE; PS00141; ASP\_PROTEASE; 1.

DR PROSITE; PS00175; ASP\_PROT\_RETROV; 1.

KW AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;

Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.

ET CHAIN 69 167

BY SIMILARITY.

ET ACT\_SITE 93 93

ET STRAND 78 82

ET STRAND 87 92

ET TURN 94 95

ET TURN 100 102

ET STRAND 111 117

ET TURN 118 119

ET STRAND 120 134

ET TURN 135 136

ET STRAND 137 146

ET STRAND 152 153

ET HELIX 155 161

ET TURN 162 162

SEQUENCE 1015 AA; 115031 MW; 164702F074A84394 CRC64;

Query Match 45.6%; Score 61.5; DB 1; Length 1015;  
 est Local Similarity 78.6%; Pred. No. 2;  
 atches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
 1 KETWETWETWTSQ 14  
 ||| |||||  
 562 KET-WETWETWTSQ 574

ULT 7  
 \_POL HVLPV STANDARD; PRT; 1015 AA.  
 P03368;  
 21-JUL-1986 (Rel. 01, Created)  
 01-NOV-1988 (Rel. 09, Last sequence update)  
 10-OCT-2003 (Rel. 42, Last annotation update)  
 POL polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.16);  
 Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].  
 POL.  
 Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).  
 Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 NCBI\_TaxID=11700;  
 [1] \_  
 SEQUENCE FROM N.A.  
 MEDLINE=8511157; PubMed=2982104;  
 Muesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,  
 Capon D.J.;  
 "Nucleic acid structure and expression of the human  
 AIDS/lymphadenopathy retrovirus";  
 Nature 313:450-458(1985).  
 [2]

REVISION.  
 Muesing M.A.;  
 Submitted (XXX-1987) to the HIV data bank.  
 -!- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,  
 and P1' variable, but often pro.  
 -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 phosphonocoeater.  
 -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate  
 + {DNA}(N).  
 -!- PTM: Cleavage sites that yield the mature proteins remain to be  
 determined.  
 -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
 KNOWN AS THE RETROPEPSIN FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 EMBL; K02083; AA059867.1; -;  
 EMBL; X01762; -; NOT\_ANNOTATED\_CDS.  
 PIR; A03967; GNVWL.  
 PDB; 1A9M; 17-JUN-98.  
 PDB; 1AJV; 20-AUG-97.  
 PDB; 1AJX; 17-SEP-97.  
 PDB; 1G35; 06-JUN-01.  
 PDB; 1GNN; 08-NOV-96.  
 PDB; 1HTG; 31-JUL-94.  
 PDB; 1HVI; 30-APR-94.  
 PDB; 1KJH; 06-MAR-02.  
 PDB; 1NPV; 04-FEB-03.  
 PDB; 1NPW; 04-FEB-03.  
 HIV; K02083; POLSFV22.  
 MEROPS; A02.001; -;  
 InterPro; IPR001969; Asparticase AS.  
 InterPro; IPR001037; Integrase C.  
 InterPro; IPR003308; Integrase Zn.  
 InterPro; IPR009007; Rept\_A acid.  
 InterPro; IPR001995; Peptidase\_A2.

DR InterPro; IPR002156; RNaseH.  
 DR InterPro; IPR001584; Rve.  
 DR InterPro; IPR000477; RVtse.  
 DR Pfam; PF00552; integrase; 1.  
 DR Pfam; PF02022; integrase; Zn; 1.  
 DR Pfam; PF00075; RNaseH; 1.  
 DR Pfam; PF00665; rve; 1.  
 DR Pfam; PF00077; rvp; 1.  
 DR Pfam; PF00078; rvt; 1.  
 DR PROSITE; PS00141; ASP\_PROTEASE; 1.  
 DR PROSITE; PS0175; ASP\_PROT\_RETROV; 1.  
 KW AIDS; Polyprotein; Hydrolyase; Aspartyl protease; Endonuclease;  
 KW Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.  
 FT CHAIN 69 167 PROTEASE  
 FT ACT\_SITE 93 93 BY SIMILARITY.  
 SQ SEQUENCE 1015 AA; 115090 MW; 51529D1BEAC2AF89 CRC64;  
 Query Match 45.6%; Score 61.5; DB 1; Length 1015;  
 Best Local Similarity 78.6%; Pred. No. 2;  
 Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 KETWETWETWTSQ 14  
 ||| |||||  
 DB 562 KET-WETWETWTSQ 574

RESULT 8  
 POL HVLPV  
 ID \_POL HVLPV STANDARD; PRT; 1006 AA.  
 AC P05961;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE POL polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.16);  
 Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].  
 GN POL.  
 OS Human immunodeficiency virus type 1 (MN isolate) (HIV-1).  
 OC Viruses; Retrovirdae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11696;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88219542; PubMed=3369091;  
 RA Gurgu C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,  
 RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;  
 RT "Envelope sequences of two new United States HIV-1 isolates";  
 RL Virology 164:531-536(1988).  
 CC -!- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,  
 and P1' variable, but often pro.  
 CC -!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 phosphonocoeater.  
 CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate  
 + {DNA}(N).  
 CC -!- PTM: Cleavage sites that yield the mature proteins remain to be  
 determined.  
 CC -!- MISCELLANEOUS: THE MN ISOLATE WAS TAKEN FROM A PEDIATRIC AIDS  
 PATIENT IN 1984.  
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO  
 KNOWN AS THE RETROPEPSIN FAMILY.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
 EMBL; M17449; -; NOT\_ANNOTATED\_CDS.  
 DR HSSP; P03366; IRVL.  
 HIV; M17449; POLSMN.  
 DR MEROPS; A02.001; -;  
 InterPro; IPR001969; Asparticase AS.  
 InterPro; IPR001037; Integrase\_C.

```

InterPro: IPR003308; Integrase Zn.
InterPro: IPR009007; Pept_A acid.
InterPro: IPR001995; Peptidase_A2.
InterPro: IPR002156; RNaseH.
InterPro: IPR001584; Rve.
InterPro: IPR000477; RVse.
Pfam: PF00552; Integrase Zn. 1.
Pfam: PF02022; Integrase Zn. 1.
Pfam: PF00075; rnaseH; 1.
Pfam: PF00665; rve; 1.
Pfam: PF00077; rvp; 1.
Pfam: PF00078; rvt; 1.
PROSITE: PS00141; ASP_PROTEASE; 1.
PROSITE: PS0175; ASP_PROT_RETROV; 1.
AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;
Nuclease; Transferase; RNA-directed DNA polymerase.
CHAIN 60 158
SITE 565 565 IN-FRAME TERMINATION CODON.
ACT SITE 84 84
SEQUENCE 1005 AA; 113860 MW; 460C50DF92AF9BB3 CRC64;

Query Match 44.8%; Score 60.5; DB 1; Length 1006;
Best Local Similarity 76.9%; Pred.No. 2.6;
Matches 10; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

1 KETWWTWTEWS 13
|||||
553 KET-WETWTEYT 564

SULT 9
POL HV1EL STANDARD; PRT; 1002 AA.
P04589; Q77906;
13-AUG-1987 (Rel. 05, Created)
01-NOV-1988 (Rel. 09, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
POL polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
P04590.
Human immunodeficiency virus type 1 (ELI isolate) (HIV-1).
Viruses; Retroviral viruses; Retroviridae; Lentivirus.
NCBI_TaxID=11689;
[1]
SEQUENCE FROM N.A.
MEDLINE=86245056; PubMed=2424612;
Alizon M., Wain-Hobson S., Montagnier L., Sonigo P.;
"Genetic variability of the AIDS virus: nucleotide sequence analysis
of two isolates from African patients.";
Cell 46:63-74(1986).
-!- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
and P1' variable, but often Pro.
-!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphonomonoester.
-!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ {DNA}(N).
-!- PTM: Cleavage sites that yield the mature proteins remain to be
determined.
-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; A07108; CAA00612.1;
EMBL; K03454; AAA44325.1; ALT_INIT.
HSPF; P03366; IHMV.
HIV; K03454; POLSELI.

```

```

DR MEROPS; A02.001; -.
DR InterPro: IPR001969; Asprotease AS.
DR InterPro: IPR001037; Integrase C.
DR InterPro: IPR003308; Integrase Zn.
DR InterPro: IPR009007; Pept_A acid.
DR InterPro: IPR001995; Peptidase_A2.
DR InterPro: IPR002156; RNaseH.
DR InterPro: IPR000477; RVse.
DR InterPro: IPR001584; Rve.
DR Pfam: PF00552; Integrase Zn. 1.
DR Pfam: PF02022; Integrase Zn. 1.
DR Pfam: PF00075; rnaseH; 1.
DR Pfam: PF00665; rve; 1.
DR Pfam: PF00077; rvp; 1.
DR Pfam: PF00078; rvt; 1.
DR PROSITE: PS00141; ASP_PROTEASE; 1.
DR PROSITE: PS0175; ASP_PROT_RETROV; 1.
AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;
KW Nuclease; Transferase; RNA-directed DNA polymerase.
FT CHAIN 56 154
FT ACT SITE 80 80 BY SIMILARITY.
FT SEQUENCE 1002 AA; 114002 MW; 570C903B689D3B54 CRC64;

Query Match 41.9%; Score 56.5; DB 1; Length 1002;
Best Local Similarity 71.4%; Pred.No. 8.1;
Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 KETWWTWTEWSQ 14
|||||
DB 549 KET-WETWAEYQ 561

RESULT 10
POL HV1A STANDARD; PRT; 1002 AA.
AC P04588; Q79582;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
GN POL.
OS Human immunodeficiency virus type 1 (MAL isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11697;
[1]
SEQUENCE FROM N.A.
MEDLINE=86245056; PubMed=2424612;
Alizon M., Wain-Hobson S., Montagnier L., Sonigo P.;
"Genetic variability of the AIDS virus: nucleotide sequence analysis
of two isolates from African patients.";
Cell 46:63-74(1986).
-!- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
and P1' variable, but often Pro.
-!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphonomonoester.
-!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ {DNA}(N).
-!- PTM: Cleavage sites that yield the mature proteins remain to be
determined.
-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X04415; CAA28012.1;
EMBL; A07116; CAA00619.1; -.

```

```

PDB; 1HHJ; 31-OCT-93.
HIV; K03456; POLSMAL.
MEROPS; A02.001; -.
InterPro; IPR001969; Asparticase AS.
InterPro; IPR001037; Integrase_C.
InterPro; IPR003308; Integrase_Zn.
InterPro; IPR009007; Pept_Acid.
InterPro; IPR001995; Peptidase_A2.
InterPro; IPR002156; RNaseH.
InterPro; IPR001584; Rve.
InterPro; IPR000477; RVtse.
Pfam; PF00552; integrase; 1.
Pfam; PF02022; Integrase_Zn; 1.
Pfam; PF00075; rnsesH; 1.
Pfam; PF00865; rve; 1.
Pfam; PF00077; rvp; 1.
Pfam; PF00078; rvt; 1.
PROSITE; PS00141; ASP PROTEASE; 1.
PROSITE; PS0175; ASP PROT RETROV; 1.
AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;
Nuclease; Transfrase; RNA-directed DNA polymerase; 3D-structure.
CHAIN 56 154
ACT SITE 80 80 BY SIMILARITY.
SEQUENCE 1002 AA; 569A29D69ACC6AC5 CRC64;

Query Match 41.9%; Score 56.5; DB 1; Length 1002;
Est Local Similarity 71.4%; Pred No. 8.1;
atches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

1 KETWETWTWTSQ 14
|||||
549 KET-WEAWTTEYMQ 561

ULT 11
_HVIRH
POL_HVIRH STANDARD; PRT; 1002 AA.
P05359;
01-NOV-1988 (Rel. 09, Created)
01-NOV-1988 (Rel. 09, Last sequence update)
28-FEB-2003 (Rel. 41, Last annotation update)
POL polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
POL.
Human immunodeficiency virus type 1 (RF/HAT isolate) (HIV-1).
Viruses; Retroviral viruses; Retroviridae; Lentivirus.
NCBI_TaxID=11701;
[1]
SEQUENCE FROM N.A.
Starcich B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S.,
Wolf H., Parks E.S., Parks W.P., Josephs S.F., Gallo R.C.,
Wong-Staal F.;
Submitted (XXX-1987) to the HIV data bank.
-!- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
and P1' variable, but often Pro.
-!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphonomoester.
-!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ [DNA](N).
-!- PTM: Cleavage sites that yield the mature proteins remain to be
determined.
-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M17451; AAA45053.1; -.

```

```

DR HSP; P04585; IRTH.
DR HIV; M17451; POLSRF.
DR MEROPS; A02.001; -.
DR InterPro; IPR001969; Asparticase AS.
DR InterPro; IPR001037; Integrase_C.
DR InterPro; IPR003308; Integrase_Zn.
DR InterPro; IPR009007; Pept_Acid.
DR InterPro; IPR001995; Peptidase_A2.
DR InterPro; IPR002156; RNaseH.
DR InterPro; IPR001584; Rve.
DR InterPro; IPR000477; RVtse.
DR Pfam; PF00552; integrase; 1.
DR Pfam; PF02022; Integrase_Zn; 1.
DR Pfam; PF00075; rnsesH; 1.
DR Pfam; PF00865; rve; 1.
DR Pfam; PF00077; rvp; 1.
DR Pfam; PF00078; rvt; 1.
DR PROSITE; PS00141; ASP PROTEASE; 1.
DR PROSITE; PS0175; ASP PROT RETROV; 1.
KW AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;
KW Nuclease; Transfrase; RNA-directed DNA polymerase.
FT CHAIN 56 154
FT ACT SITE 80 80 BY SIMILARITY.
SQ SEQUENCE 1002 AA; 864341718E5C48C2 CRC64;

Query Match 41.9%; Score 56.5; DB 1; Length 1002;
Best Local Similarity 71.4%; Pred No. 8.1;
Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 KETWETWTWTSQ 14
|||||
DB 549 KET-WEAWTTEYMQ 561

RESULT 12
POL_HV1Z2 STANDARD; PRT; 1002 AA.
ID POL_HV1Z2
AC P12499;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE POL polyprotein [Contains: Protease (Retropepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
POL.
Human immunodeficiency virus type 1 (Z2/CDC-234 isolate) (HIV-1).
Viruses; Retroviral viruses; Retroviridae; Lentivirus.
NCBI_TaxID=11683;
[1]
SEQUENCE FROM N.A.
Theodore T., Buckler-White A.;
Submitted (NOV-1988) to the HIV data bank.
-!- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
and P1' variable, but often Pro.
-!- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphonomoester.
-!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ [DNA](N).
-!- PTM: Cleavage sites that yield the mature proteins remain to be
determined.
-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO
KNOWN AS THE RETROPEPSIN FAMILY.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; M22639; AAA45366.1; -.
PIR; S54378; S54378.
PDB; 1E28; 12-SEP-00.

```

```

PDB; 1HXV; 04-FEB-98.
PDB; 1PRO; 17-AUG-96.
PDB; 1V1J; 13-JAN-99.
PDB; 1V1K; 13-JAN-99.
HIV; M22639; POLS2226.
MEROPS; A02.001; -.
InterPro; IPR001969; Asparticase AS.
InterPro; IPR001037; Integrase C.
InterPro; IPR003308; Integrase Zn.
InterPro; IPR009007; Pept_A_acid.
InterPro; IPR001995; Peptidase_A2.
InterPro; IPR002156; RNaseH.
InterPro; IPR001584; Rve.
InterPro; IPR000477; RVTse.
Pfam; PF00582; Integrase_1.
Pfam; PF02022; Integrase_Zn; 1.
Pfam; PF00075; RnaseH; 1.
Pfam; PF00077; Rvp; 1.
Pfam; PF00078; Rvt; 1.
PROSITE; PS00141; ASP_PROTEASE; 1.
PROSITE; PS00175; ASP_PROT_RETROV; 1.
AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;
Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
CHAIN 56 154
ACT SITE 80 80 BY SIMILARITY.
SEQUENCE 1002 AA; 113724 MW; CB4AAC9AB4742315 CRC64;

Query Match 41.9%; Score 56.5; DB 1; Length 1002;
Best Local Similarity 71.4%; Pred. NO. 8.1;
Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

1 KETWETWTEWSQ 14
||| ||||| |
549 KET-WETWVEYVQ 561

SULT 13
L-HVINS
POL_HVINS STANDARD; PRT; 1003 AA.
P12497;
01-OCT-1989 (Rel. 12, Created)
01-OCT-1989 (Rel. 12, Last sequence update)
10-OCT-2003 (Rel. 42, Last annotation update)
POL polyprotein [Contains: Protease (Retropoepsin) (EC 3.4.23.16);
Reverse transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)].
POL.
Human immunodeficiency virus type 1 (New York-5 isolate) (HIV-1).
Viruses; Retroviral viruses; Retroviridae; Lentivirus.
NCBI_TaxID=11698;
[1]
SEQUENCE FROM N.A. (CLONE PNL4-3).
Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;
Submitted (JUN-1988) to the EMBL/GenBank/DBJ databases.
[2]
X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 57-155.
MEDLINE=90354401; PubMed=2201682;
Fitzgerald P.M.D., McKeever B.M., van Middlesworth J.F.,
Springer J.P., Heinbach J.C., Leu C.-T., Herber W.K., Dixon R.A.F.,
Darke P.L.;
"Crystallographic analysis of a complex between human
immunodeficiency virus type 1 protease and acetyl-pepstatin at 2.0-A
resolution.";
J. Biol. Chem. 265:14209-14219(1990).
-1- CATALYTIC ACTIVITY: Specific for a P1 residue that is hydrophobic,
and P1' variable, but often pro.
-1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
phosphonooester.
-1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
+ {DNA}(N).
-1- PTM: Cleavage sites that yield the mature proteins remain to be
determined.
-1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY A2; ALSO

```

```

CC CC KNOWN AS THE RETROPEPSIN FAMILY.
CC CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC CC the European Bioinformatics Institute. There are no restrictions on its
CC CC use by non-profit institutions as long as its content is in no way
CC CC modified and this statement is not removed. Usage by and for commercial
CC CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC CC or send an email to license@isb-sib.ch).
CC CC -----
CC CC EMBL; M19921; AAA44988.1; -.
CC CC PDB; 5HVP; 15-OCT-91.
CC CC PDB; 4PHV; 31-OCT-93.
CC CC PDB; 1B3F; 19-JUL-99.
CC CC PDB; 1BHL; 28-OCT-98.
CC CC PDB; 1B14; 18-NOV-98.
CC CC PDB; 1BIS; 16-SEP-98.
CC CC PDB; 1BIU; 16-SEP-98.
CC CC PDB; 1WJD; 13-MAY-98.
CC CC PDB; 1WJD; 13-MAY-98.
CC CC PDB; 2ITG; 12-MAR-97.
CC CC PDB; 9HVP; 15-JUL-92.
CC CC HIV; M19921; POLS43.
CC CC MEROPS; A02.001; -.
CC CC InterPro; IPR001969; Asparticase AS.
CC CC InterPro; IPR001037; Integrase C.
CC CC InterPro; IPR003308; Integrase Zn.
CC CC InterPro; IPR009007; Pept_A_acid.
CC CC InterPro; IPR001995; Peptidase_A2.
CC CC InterPro; IPR002156; RNaseH.
CC CC InterPro; IPR001584; Rve.
CC CC InterPro; IPR000477; RVTse.
CC CC Pfam; PF00582; Integrase_1.
CC CC Pfam; PF02022; Integrase_Zn; 1.
CC CC Pfam; PF00075; RnaseH; 1.
CC CC Pfam; PF00077; Rvp; 1.
CC CC Pfam; PF00078; Rvt; 1.
CC CC PROSITE; PS00141; ASP_PROTEASE; 1.
CC CC PROSITE; PS00175; ASP_PROT_RETROV; 1.
CC CC AIDS; Polyprotein; Hydrolase; Aspartyl protease; Endonuclease;
CC CC Nuclease; Transferase; RNA-directed DNA polymerase; 3D-structure.
CC CC CHAIN 57 155
CC CC ACT SITE 81 81
CC CC FT STRAND 58 60
CC CC FT STRAND 66 71
CC CC FT TURN 72 73
CC CC FT STRAND 74 80
CC CC FT TURN 82 83
CC CC FT STRAND 87 90
CC CC FT STRAND 99 105
CC CC FT TURN 106 107
CC CC FT STRAND 108 122
CC CC FT TURN 123 124
CC CC FT STRAND 125 134
CC CC FT STRAND 140 141
CC CC FT HELIX 143 146
CC CC FT HELIX 147 149
CC CC FT TURN 150 150
CC CC FT STRAND 152 154
CC CC SQ SEQUENCE 1003 AA; 113535 MW; EB9CDBA30E114A55 CRC64;

Query Match 41.9%; Score 56.5; DB 1; Length 1003;
Best Local Similarity 71.4%; Pred. NO. 8.1;
Matches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 KETWETWTEWSQ 14
||| ||||| |
Db 550 KET-WEAWWTEYVQ 562

RESULT 14
POL_HVIOY

```







SULT 2  
TVU1 Q8TVU1 PRELIMINARY; PRT; 793 AA.  
Q8TVU1: 01-JUN-2002 (T-REMBLrel. 21, Created)  
01-JUN-2002 (T-REMBLrel. 21, Last sequence update)  
01-JUN-2003 (T-REMBLrel. 24, Last annotation update)  
Predicted DNA-dependent DNA polymerase, component of a  
thermophile-specific DNA repair system.  
MK1297.  
Methanopyrus kandleri.  
Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;  
Methanopyrus.  
NCBI\_TaxID=2320;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=AV19 / DSM 6324 / JCM 9639;  
MEDLINE=21927647; PubMed=11930014;  
Shcherbina A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,  
Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,  
Natalie D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,  
Malykh A.G., Koonin E.V., Kozlyavkin S.A.,  
"The complete genome of hyperthermophile Methanopyrus kandleri AV19  
and monophyly of archaeal methanogens.";  
Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649 (2002).  
EMBL; AE010421; AA002510.1; -  
InterPro; IPR000160; GGDFF.  
PROSITE; PS00887; GGDFF; 1.  
Complete proteome.  
SEQUENCE 793 AA; 88610 MW; 1954F3B9975A8F63 CRC64;  
Query Match 46.7%; Score 63; DB 17; Length 793;  
Best Local Similarity 61.5%; Pred. No. 4.1;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
4 WWTWWTWWSQPK 16  
711 WWTWWTWWSQPK 723  
POL.  
Human immunodeficiency virus 1.  
Viruses; Retroviridae; Retroviridae; Lentivirus.  
NCBI\_TaxID=11676;  
[1]  
SEQUENCE FROM N.A.  
STRAIN=19142-16-polB;  
MEDLINE=99214336; PubMed=10196293;  
Wilson C.C., Brown R.C., Korber B.T., Wilkes B.M., Ruhl D.J.,  
Sakamoto D., Kunstman K., Luzuriaga K., Hanson I.C., Widmayer S.M.,  
Wiznia A., Clapp S., Ammann A.J., Koup R.A., Wolinsky S.M.,  
Walker B.D.;  
"Frequent detection of escape from cytotoxic T-lymphocyte recognition  
in perinatal human immunodeficiency virus (HIV) type 1 transmission:  
the arxel project for the prevention of transmission of HIV from  
mother to infant.";  
J. Virol. 73:3975-3985 (1999).  
EMBL; AF212616; AA28988.1;  
GO; GO:0003676; F:nucleic acid binding; IEA.

DR GO; GO:0004523; F:ribonuclease H activity; IEA.  
DR InterPro; IPR002156; RNaseH.  
DR Pfam; PF00075; RNaseH; 1.  
FT NON\_TER 1  
SQ SEQUENCE 212 AA; 24550 MW; F7756F76D5298EDB CRC64;  
Query Match 45.6%; Score 61.5; DB 15; Length 212;  
Best Local Similarity 78.6%; Pred. No. 1.8;  
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
QY 1 KETWETWWTWWSQ 14  
DB 121 KETWETWWTWWSQ 133  
RESULT 4  
Q998V3 PRELIMINARY; PRT; 225 AA.  
AC Q998V3;  
DT 01-JUN-2001 (T-REMBLrel. 17, Created)  
DT 01-JUN-2001 (T-REMBLrel. 17, Last sequence update)  
DT 01-OCT-2003 (T-REMBLrel. 25, Last annotation update)  
DE Pol protein (Fragment).  
GN POL.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=97DC.KFE45;  
RX MEDLINE=20584657; PubMed=11153090;  
RA Vidal N., Mulara-Kabeya C., Nzilambi N., Delaporte E., Peeters M.;  
"Identification of a complex env subtype E HIV type 1 virus from the  
Democratic Republic of Congo, recombinant with A, G, H, J, K, and  
unknown subtypes.";  
RT AIDS Res. Hum. Retroviruses 16:2059-2064 (2000).  
RL ENBL; A404322; CAC24829.1; -  
DR HSP; P03366; IHRH.  
DR GO; GO:0003676; F:nucleic acid binding; IEA.  
DR GO; GO:0004523; F:ribonuclease H activity; IEA.  
DR InterPro; IPR002156; RNaseH.  
DR Pfam; PF00075; RNaseH; 1.  
FT NON\_TER 1  
FT NON\_TER 225  
SQ SEQUENCE 225 AA; 25904 MW; 3EC5744F6FADAF46 CRC64;  
Query Match 45.6%; Score 61.5; DB 15; Length 225;  
Best Local Similarity 78.6%; Pred. No. 1.9;  
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
QY 1 KETWETWWTWWSQ 14  
DB 66 KETWETWWTWWSQ 78  
RESULT 5  
Q9WGX3 PRELIMINARY; PRT; 236 AA.  
AC Q9WGX3;  
DT 01-NOV-1999 (T-REMBLrel. 12, Created)  
DT 01-NOV-1999 (T-REMBLrel. 12, Last sequence update)  
DT 01-OCT-2003 (T-REMBLrel. 25, Last annotation update)  
DE Pol protein (Fragment).  
GN POL.  
OS Human immunodeficiency virus 1.  
OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=08106-C6A-5-polB;  
RX MEDLINE=99214336; PubMed=10196293;  
RA Wilson C.C., Brown R.C., Korber B.T., Wilkes B.M., Ruhl D.J.,  
Sakamoto D., Kunstman K., Luzuriaga K., Hanson I.C., Widmayer S.M.,

Wiznia A., Clapp S., Ammann A.J., Koup R.A., Wolinsky S.M., Walker B.D.;  
 "Frequent detection of escape from cytotoxic T-lymphocyte recognition in perinatal human immunodeficiency virus (HIV) type 1 transmission: the ariel project for the prevention of transmission of HIV from mother to infant.";  
 J. Virol. 73:3975-3985(1999).  
 EMBL; AF121647; AAD29019.1; -.  
 HSSP; P03366; 1HRH.  
 GO; GO:0003676; F:nucleic acid binding; IEA.  
 GO; GO:0004523; F:ribonuclease H activity; IEA.  
 InterPro; IPR002156; RNaseH.  
 Pfam; PF00075; rnaseh; 1.  
 NON\_TER 1  
 NON\_TER 236  
 SEQUENCE 236 AA; 26971 MW; 797EFD34F3DB01A9 CRC64;  
 uery Match 45.6%; Score 61.5; DB 15; Length 236;  
 est Local Similarity 71.4%; Pred. No. 2;  
 atches 10; Conservative 2; Mismatches 1; Indels 1; Gaps 1;  
 1 KETWETWWTWSQ 14  
 ||| |||||:  
 120 KET-WETWWTGYQ 132

ULT 6

GVO Q9WGV0 PRELIMINARY; PRT; 237 AA.  
 Q9WGV0;  
 01-NOV-1999 (TrEMBLrel. 12, Created)  
 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 Polyprotein (Fragment).  
 POL.  
 Human immunodeficiency virus 1.  
 Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
 NCBI\_TaxID=11676;  
 [1] \_  
 SEQUENCE FROM N.A.  
 STRAIN=19142-6-polB;  
 MEDLINE=9214336; PubMed=10196293;  
 Wilson C.C., Brown R.C., Korber B.T., Wilkes B.M., Ruhl D.J., Sakamoto D., Kunstman K., Luzuriaga K., Hanson I.C., Widmayer S.M., Wiznia A., Clapp S., Ammann A.J., Koup R.A., Wolinsky S.M., Walker B.D.;  
 "Frequent detection of escape from cytotoxic T-lymphocyte recognition in perinatal human immunodeficiency virus (HIV) type 1 transmission: the ariel project for the prevention of transmission of HIV from mother to infant.";  
 J. Virol. 73:3975-3985(1999).  
 EMBL; AF121621; AAD28993.1; -.  
 HSSP; P03366; 1HRH.  
 GO; GO:0003676; F:nucleic acid binding; IEA.  
 GO; GO:0004523; F:ribonuclease H activity; IEA.  
 InterPro; IPR002156; RNaseH.  
 Pfam; PF00075; rnaseh; 1.  
 NON\_TER 1  
 NON\_TER 237  
 SEQUENCE 237 AA; 27108 MW; BF8991B93E95AEFB CRC64;  
 uery Match 45.6%; Score 61.5; DB 15; Length 237;  
 est Local Similarity 78.6%; Pred. No. 2;  
 atches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
 1 KETWETWWTWSQ 14  
 ||| |||||:  
 121 KET-WETWWTGYQ 133

ULT 7

GVS Q9WGV3 PRELIMINARY; PRT; 237 AA.

AC Q9WGV3;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Polyprotein (Fragment).  
 GN POL.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1] \_  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=08102-A4-polB;  
 RX MEDLINE=9214336; PubMed=10196293;  
 RA Wilson C.C., Brown R.C., Korber B.T., Wilkes B.M., Ruhl D.J., Sakamoto D., Kunstman K., Luzuriaga K., Hanson I.C., Widmayer S.M., Wiznia A., Clapp S., Ammann A.J., Koup R.A., Wolinsky S.M., Walker B.D.;  
 RA "Frequent detection of escape from cytotoxic T-lymphocyte recognition in perinatal human immunodeficiency virus (HIV) type 1 transmission: the ariel project for the prevention of transmission of HIV from mother to infant.";  
 RL J. Virol. 73:3975-3985(1999).  
 DR EMBL; AF121624; AAD28996.1; -.  
 DR HSSP; P03366; 1HRH.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.  
 DR InterPro; IPR002156; RNaseH.  
 DR Pfam; PF00075; rnaseh; 1.  
 FT NON\_TER 1  
 FT NON\_TER 237  
 SQ SEQUENCE 237 AA; 27449 MW; A60D9EF956D8C51 CRC64;  
 Query Match 45.6%; Score 61.5; DB 15; Length 237;  
 Best Local Similarity 78.6%; Pred. No. 2;  
 Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
 QY 1 KETWETWWTWSQ 14  
 ||| |||||:  
 DB 121 KET-WETWWTGYQ 133

RESULT 8  
 Q9WGV7  
 ID Q9WGV7 PRELIMINARY; PRT; 237 AA.  
 AC Q9WGV7;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Polyprotein (Fragment).  
 GN POL.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retrovird viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1] \_  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=08102-G3-polB;  
 RX MEDLINE=9214336; PubMed=10196293;  
 RA Wilson C.C., Brown R.C., Korber B.T., Wilkes B.M., Ruhl D.J., Sakamoto D., Kunstman K., Luzuriaga K., Hanson I.C., Widmayer S.M., Wiznia A., Clapp S., Ammann A.J., Koup R.A., Wolinsky S.M., Walker B.D.;  
 RA "Frequent detection of escape from cytotoxic T-lymphocyte recognition in perinatal human immunodeficiency virus (HIV) type 1 transmission: the ariel project for the prevention of transmission of HIV from mother to infant.";  
 RL J. Virol. 73:3975-3985(1999).  
 DR EMBL; AF121628; AAD29000.1; -.  
 DR HSSP; P03366; 1HRH.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.  
 DR InterPro; IPR002156; RNaseH.  
 DR Pfam; PF00075; rnaseh; 1.  
 FT NON\_TER 1

```

NON TER      237      237
SEQUENCE     237 AA; 27384 MW; F0E6D1A729FDD3C3 CRC64;

Query Match      45.6%; Score 61.5; DB 15; Length 237;
Best Local Similarity 78.6%; Pred. No. 2;
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

      1 KETWETWWTWSQ 14
      ||| ||||| |||
      121 KET-WETWWTWYQ 133

SULT 9
QWGW5
QWGW5 PRELIMINARY; PRT; 237 AA.
01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Polyprotein (Fragment).
POL.
Human immunodeficiency virus 1.
Viruses; Retrovirdae; Retroviridae; Lentivirus.
NCBI_TaxID=11676;
[1]
SEQUENCE FROM N.A.
STRAIN=02113-B2-6-polB;
MEDLINE=99214336; PubMed=10196293;
Wilson C.C., Brown R.C., Korber B.T., Wilkes B.M., Ruhl D.J.,
Sakamoto D., Kunstman K., Luzuriaga K., Hanson I.C., Widmayer S.M.,
Wiznia A., Clapp S., Ammann A.J., Koup R.A., Wolinsky S.M.,
Walker B.D.;
"Frequent detection of escape from cytotoxic T-lymphocyte recognition
in perinatal human immunodeficiency virus (HIV) type 1 transmission:
the ariel project for the prevention of transmission of HIV from
mother to infant."
J. Virol. 73:3975-3985(1999).
EMBL; AF121636; AAD29008.1; -.
HSP; P03366; IHRH.
GO; GO:0003676; F:nucleic acid binding; IEA.
InterPro: IPR002156; RNaseH.
Pfam; PF00075; rnaaseh; 1.
NON TER      237      237
SEQUENCE     237 AA; 27236 MW; 50FEB2CD2A2739C6 CRC64;

Query Match      45.6%; Score 61.5; DB 15; Length 237;
Best Local Similarity 78.6%; Pred. No. 2;
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

      1 KETWETWWTWSQ 14
      ||| ||||| |||
      121 KET-WETWWTWYQ 133

SULT 10
QWGW0
QWGW0 PRELIMINARY; PRT; 237 AA.
01-NOV-1999 (TrEMBLrel. 12, Created)
01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
Polyprotein (Fragment).
POL.
Human immunodeficiency virus 1.
Viruses; Retrovirdae; Retroviridae; Lentivirus.
NCBI_TaxID=11676;
[1]
SEQUENCE FROM N.A.
STRAIN=08102-M-1-polB;
MEDLINE=99214336; PubMed=10196293;
Wilson C.C., Brown R.C., Korber B.T., Wilkes B.M., Ruhl D.J.,

```

```

RA Sakamoto D., Kunstman K., Luzuriaga K., Hanson I.C., Widmayer S.M.,
RA Wiznia A., Clapp S., Ammann A.J., Koup R.A., Wolinsky S.M.,
RA Walker B.D.;
RT "Frequent detection of escape from cytotoxic T-lymphocyte recognition
RT in perinatal human immunodeficiency virus (HIV) type 1 transmission:
RT the ariel project for the prevention of transmission of HIV from
RT mother to infant."
RL J. Virol. 73:3975-3985(1999).
DR EMBL; AF121631; AAD29003.1; -.
DR HSP; P03366; IHRH.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR InterPro: IPR002156; RNaseH.
DR Pfam; PF00075; rnaaseh; 1.
FT NON_TER      237      237
FT NON_TER      237      237
SQ SEQUENCE     237 AA; 27457 MW; 3F521B7E5B1C75AF CRC64;

Query Match      45.6%; Score 61.5; DB 15; Length 237;
Best Local Similarity 78.6%; Pred. No. 2;
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY      1 KETWETWWTWSQ 14
DB      ||| ||||| |||
      121 KET-WETWWTWYQ 133

RESULT 11
QWGW5
ID QWGW5 PRELIMINARY; PRT; 237 AA.
AC QWGW5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Polyprotein (Fragment).
DE POL.
GN POL.
OS Human immunodeficiency virus 1.
OC Viruses; Retrovirdae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=02113-B2-3-polB;
RC MEDLINE=99214336; PubMed=10196293;
EX Wilson C.C., Brown R.C., Korber B.T., Wilkes B.M., Ruhl D.J.,
RA Sakamoto D., Kunstman K., Luzuriaga K., Hanson I.C., Widmayer S.M.,
RA Wiznia A., Clapp S., Ammann A.J., Koup R.A., Wolinsky S.M.,
RA Walker B.D.;
RA "Frequent detection of escape from cytotoxic T-lymphocyte recognition
RT in perinatal human immunodeficiency virus (HIV) type 1 transmission:
RT the ariel project for the prevention of transmission of HIV from
RT mother to infant."
RL J. Virol. 73:3975-3985(1999).
DR EMBL; AF121626; AAD28998.1; -.
DR HSP; P03366; IHRH.
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR GO; GO:0004523; F:ribonuclease H activity; IEA.
DR InterPro: IPR002156; RNaseH.
DR Pfam; PF00075; rnaaseh; 1.
FT NON_TER      237      237
FT NON_TER      237      237
SQ SEQUENCE     237 AA; 27262 MW; 4A5BF3C5E5B2FA7 CRC64;

Query Match      45.6%; Score 61.5; DB 15; Length 237;
Best Local Similarity 78.6%; Pred. No. 2;
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY      1 KETWETWWTWSQ 14
DB      ||| ||||| |||
      121 KET-WETWWTWYQ 133

RESULT 12
QWGW9

```

Q9WGV9 PRELIMINARY; PRT; 237 AA.  
 Q9WGV9; STRAIN=08102-L-3-POLB;  
 01-NOV-1999 (TREMELrel. 12, Created)  
 01-NOV-1999 (TREMELrel. 12, Last sequence update)  
 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 Polyprotein (Fragment).  
 POL.  
 Human immunodeficiency virus 1.  
 Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 NCBI\_TaxID=11676;  
 [1]  
 SEQUENCE FROM N.A.  
 STRAIN=08102-L-3-POLB;  
 MEDLINE=99214336; PubMed=10196293;  
 Wilson C.C., Brown R.C., Korber B.T., Wilkes B.M., Ruhl D.J.,  
 Sakamoto D., Kunstman K., Luzuriaga K., Hanson I.C., Widmayer S.M.,  
 Wiznia A., Clapp S., Ammann A.J., Koup R.A., Wolinsky S.M.,  
 Walker B.D.;  
 "Frequent detection of escape from cytotoxic T-lymphocyte recognition  
 in perinatal human immunodeficiency virus (HIV) type 1 transmission:  
 the ariel project for the prevention of transmission of HIV from  
 mother to infant.";  
 J. Virol. 73:3975-3985(1999).  
 EMBL; AF121630; AAD29002.1; -.  
 HSSP; P03366; IHRH.  
 GO; GO:0003676; F:nucleic acid binding; IEA.  
 GO; GO:0004523; F:ribonuclease H activity; IEA.  
 InterPro; IPR002156; RNaseH.  
 Pfam; PF00075; rnaseh; 1.  
 NON\_TER 1  
 NON\_TER 237  
 SEQUENCE 237 AA; 27381 MW; A9411F47E8E8D0E93 CRC64;  
 Query Match 45.6%; Score 61.5; DB 15; Length 237;  
 Best Local Similarity 78.6%; Pred. No. 2;  
 Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
 [1]  
 1 KETWETWTEWSQ 14  
 ||| |||||:  
 121 KET-WETWTEWYQ 133

ULT 13  
 48H5

Q9WH5 PRELIMINARY; PRT; 237 AA.  
 Q9WH5; STRAIN=08102-L-3-POLB;  
 01-NOV-1999 (TREMELrel. 12, Created)  
 01-NOV-1999 (TREMELrel. 12, Last sequence update)  
 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 Polyprotein (Fragment).  
 POL.  
 Human immunodeficiency virus 1.  
 Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 NCBI\_TaxID=11676;  
 [1]  
 SEQUENCE FROM N.A.  
 STRAIN=08102-L-3-POLB;  
 MEDLINE=99214336; PubMed=10196293;  
 Wilson C.C., Brown R.C., Korber B.T., Wilkes B.M., Ruhl D.J.,  
 Sakamoto D., Kunstman K., Luzuriaga K., Hanson I.C., Widmayer S.M.,  
 Wiznia A., Clapp S., Ammann A.J., Koup R.A., Wolinsky S.M.,  
 Walker B.D.;  
 "Frequent detection of escape from cytotoxic T-lymphocyte recognition  
 in perinatal human immunodeficiency virus (HIV) type 1 transmission:  
 the ariel project for the prevention of transmission of HIV from  
 mother to infant.";  
 J. Virol. 73:3975-3985(1999).  
 EMBL; AF121630; AAD29002.1; -.  
 HSSP; P03366; IHRH.  
 GO; GO:0003676; F:nucleic acid binding; IEA.  
 GO; GO:0004523; F:ribonuclease H activity; IEA.

DR InterPro; IPR002156; RNaseH.  
 DR Pfam; PF00075; rnaseh; 1.  
 KW Polyprotein.  
 FT NON\_TER 1  
 FT NON\_TER 237  
 SQ SEQUENCE 237 AA; 27321 MW; 3EBCE634B25333 CRC64;  
 Query Match 45.6%; Score 61.5; DB 15; Length 237;  
 Best Local Similarity 78.6%; Pred. No. 2;  
 Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
 [1]  
 1 KETWETWTEWSQ 14  
 ||| |||||:  
 121 KET-WETWTEWYQ 133

DB 121 KET-WETWTEWYQ 133

RESULT 14  
 Q9W9UO PRELIMINARY; PRT; 237 AA.  
 ID Q9W9UO; STRAIN=19142-3-POLB, and 19142-20-POLB;  
 AC Q9W9UO; MEDLINE=99214336; PubMed=10196293;  
 DT 01-NOV-1999 (TREMELrel. 12, Created)  
 DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Polyprotein (Fragment).  
 GN POL.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=19142-3-POLB, and 19142-20-POLB;  
 RX MEDLINE=99214336; PubMed=10196293;  
 RA Wilson C.C., Brown R.C., Korber B.T., Wilkes B.M., Ruhl D.J.,  
 Sakamoto D., Kunstman K., Luzuriaga K., Hanson I.C., Widmayer S.M.,  
 Wiznia A., Clapp S., Ammann A.J., Koup R.A., Wolinsky S.M.,  
 Walker B.D.;  
 RA "Frequent detection of escape from cytotoxic T-lymphocyte recognition  
 in perinatal human immunodeficiency virus (HIV) type 1 transmission:  
 the ariel project for the prevention of transmission of HIV from  
 mother to infant.";  
 RT J. Virol. 73:3975-3985(1999).  
 RL EMBL; AF121620; AAD28992.1; -.  
 DR EMBL; AF121618; AAD28990.1; -.  
 DR HSSP; P03366; IHRH.  
 DR GO; GO:0003676; F:nucleic acid binding; IEA.  
 DR GO; GO:0004523; F:ribonuclease H activity; IEA.  
 DR InterPro; IPR002156; RNaseH.  
 DR Pfam; PF00075; rnaseh; 1.  
 KW Polyprotein.  
 FT NON\_TER 1  
 FT NON\_TER 237  
 SQ SEQUENCE 237 AA; 27208 MW; B26BE9624E54F7CA CRC64;  
 Query Match 45.6%; Score 61.5; DB 15; Length 237;  
 Best Local Similarity 78.6%; Pred. No. 2;  
 Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
 [1]  
 1 KETWETWTEWSQ 14  
 ||| |||||:  
 121 KET-WETWTEWYQ 133

DB 121 KET-WETWTEWYQ 133

RESULT 15  
 Q9WGX6 PRELIMINARY; PRT; 237 AA.  
 ID Q9WGX6; STRAIN=08102-L-3-POLB;  
 AC Q9WGX6; MEDLINE=99214336; PubMed=10196293;  
 DT 01-NOV-1999 (TREMELrel. 12, Created)  
 DT 01-NOV-1999 (TREMELrel. 12, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Polyprotein (Fragment).  
 GN POL.  
 OS Human immunodeficiency virus 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

```

NCBI_TaxID=41676;
[1]
SEQUENCE FROM N.A.
STRAIN=02113-B7B-2-polB;
MEDLINE=9214336; PubMed=10196293;
Wilson C.C., Brown R.C., Korber B.T., Wilkes B.M., Ruhl D.J.,
Sakamoto D., Kunstman K., Luzuriaga K., Hanson I.C., Widmayer S.M.,
Wiznia A., Clapp S., Ammann A.J., Koup R.A., Wolinsky S.M.,
Walker B.D.;
"Frequency detection of escape from cytotoxic T-lymphocyte recognition
in perinatal human immunodeficiency virus (HIV) type 1 transmission:
the ariel project for the prevention of transmission of HIV from
mother to infant.";
J. Virol. 73:3975-3985(1999).
EMBL; AF121652; AAD29024.1; -.
HSSP; P03366; 1HRH.
GO; GO:0003676; F:nucleic acid binding; IEA.
GO; GO:0004523; F:ribonuclease H activity; IEA.
InterPro; IPR002156; RNaseH.
Pfam; PF00075; RNaseH; 1.
NON_TER 1
NON_TER 237 237
SEQUENCE 237 AA; 27335 MW; BAFCA15760BDAA73 CRC64;

Query Match 45.6%; Score 61.5; DB 15; Length 237;
Best Local Similarity 78.6%; Pred No. 2;
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

1 KETWETWTEWSQ 14
||| |||||
121 KET-WETWTEYMQ 133

```

arch completed: February 25, 2004, 14:48:15  
 b time : 43 secs



TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1

FEATURE:  
NAME/KEY: VARIANT  
LOCATION: 100  
OTHER INFORMATION: Xaa=Leu or Ile

FEATURE:  
NAME/KEY: VARIANT  
LOCATION: 103  
OTHER INFORMATION: Xaa=Lys or Asn

FEATURE:  
NAME/KEY: VARIANT  
LOCATION: 106  
OTHER INFORMATION: Xaa=Val, Ile or Ala

FEATURE:  
NAME/KEY: VARIANT  
LOCATION: 108  
OTHER INFORMATION: Xaa=Val or Ile

FEATURE:  
NAME/KEY: VARIANT  
LOCATION: 138  
OTHER INFORMATION: Xaa=Glu or Lys

FEATURE:  
NAME/KEY: VARIANT  
LOCATION: 181  
OTHER INFORMATION: Xaa=Tyr or Cys

FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (0)...(0)  
OTHER INFORMATION: Xaa at position 188 can be Tyr or Cys

FEATURE:  
NAME/KEY: VARIANT  
LOCATION: (0)...(0)  
OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

OTHER INFORMATION: Xaa at position 236 can be Pro or Leu

Db 395 KET-WETWTEYQ 407

RESULT 4

US-09-735-487-14

; Sequence 14, Application US/09735487

; Patent No. 6528251

; GENERAL INFORMATION:

; APPLICANT: de BETHUNE, Marie-Pierre

; APPLICANT: HERTOOGS, Kurt

; APPLICANT: PAUWELS, Rudi

; TITLE OF INVENTION: METHOD OF MANAGING THE CHEMOTHERAPY OF PATIENTS WHO ARE

; TITLE OF INVENTION: HIV POSITIVE BASED ON THE PHENOTYPIC DRUG SENSITIVITY

; TITLE OF INVENTION: OF HUMAN HIV STRAINS

; FILE REFERENCE: 1377-125P

; CURRENT APPLICATION NUMBER: US/09/735,487

; CURRENT FILING DATE: 2000-12-14

; PRIOR APPLICATION NUMBER: 09/117,217

; PRIOR FILING DATE: 1998-07-24

; NUMBER OF SEQ ID NOS: 15

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 14

; LENGTH: 562

; TYPE: PRT

; ORGANISM: HIV-HXB2

; US-09-735-487-14

Query Match 45.6%; Score 61.5; DB 4; Length 562;

Best Local Similarity 78.6%; Pred. No. 1.7;

Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 KETWETWTEYQ 14

Db 395 KET-WETWTEYQ 407

RESULT 5

US-07-743-357-22

; Sequence 22, Application US/07743357

; Patent No. 5858646

; GENERAL INFORMATION:

; APPLICANT: Kang, Yong C.

; TITLE OF INVENTION: Polypeptide having immunological

; TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: KIRBY EADES GALE BAKER

; STREET: Box 3432, Station D

; CITY: Ottawa

; STATE: Ontario

; COUNTRY: Canada

; ZIP: KIM 1H6

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/743,357

; FILING DATE: 21-AUG-1991

; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: PCT/CA90/00062

; FILING DATE: 23-FEB-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Gale, Edwin J.

; REGISTRATION NUMBER: 28,584

; REFERENCE/DOCKET NUMBER: 30924-2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (613) 237-6900

; TELEFAX: (613) 237-0045

; INFORMATION FOR SEQ ID NO: 22:

; SEQUENCE CHARACTERISTICS:



LENGTH: 913 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Human immunodeficiency virus type 1  
STRAIN: HXB2  
07-743-357-22

very Match 45.6%; Score 61.5; DB 2; Length 913;  
est Local Similarity 78.6%; Pred. No. 2.8;  
atches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

1 KETWETWTEWSQ 14  
||| |||||:  
459 KET-WETWTEYQ 471

JLT 6  
07-743-357-1  
sequence 1, Application US/07743357  
atent No. 5858646  
GENERAL INFORMATION:  
APPLICANT: Kang, Yong C.  
TITLE OF INVENTION: Polypeptide having immunological  
TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KIRBY EADES GALE BAKER  
STREET: Box 3432, Station D  
CITY: Ottawa  
STATE: Ontario  
COUNTRY: Canada  
ZIP: K1M 1H8

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/743,357  
FILING DATE: 21-AUG-1991  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/CA90/00062  
FILING DATE: 23-FEB-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Gale, Edwin J.  
REGISTRATION NUMBER: 28,584  
REFERENCE/DOCKET NUMBER: 30924-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (613) 237-6900  
TELEFAX: (613) 237-0045  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1005 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Human immunodeficiency virus type 1  
STRAIN: HXB2  
07-743-357-1

very Match 45.6%; Score 61.5; DB 2; Length 1005;  
est Local Similarity 78.6%; Pred. No. 3.1;  
atches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 1 KETWETWTEWSQ 14  
||| |||||:  
Db 551 KET-WETWTEYQ 563

RESULT 7  
US-08-463-210-9  
; Sequence 9, Application US/08463210  
; Patent No. 6001977  
; GENERAL INFORMATION:  
; APPLICANT: CHANG, Nancy T.  
; APPLICANT: GALLO, Robert C.  
; APPLICANT: WONG-STALL, Flossie  
; TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morgan & Finnegan, L.L.P.  
; STREET: 345 Park Avenue  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10154-0053  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/463,210  
; FILING DATE: 05-JUN-1995  
; CLASSIFICATION: 436  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 06/693,866  
; FILING DATE: 23-JAN-1985  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 06/659,339  
; FILING DATE: 10-OCT-1984  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Serunian, Leslie A.  
; REGISTRATION NUMBER: 35,353  
; REFERENCE/DOCKET NUMBER: 2026-4193US2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 758-4800  
; TELEFAX: (212) 751-6849  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1015 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; ORGANISM: HTLV-III  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..1015  
; OTHER INFORMATION: /note= "pol protein of HTLV-III"  
US-08-463-210-9

Query Match 45.6%; Score 61.5; DB 3; Length 1015;  
Best Local Similarity 78.6%; Pred. No. 3.1;  
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Qy 1 KETWETWTEWSQ 14  
||| |||||:  
Db 562 KET-WETWTEYQ 574

RESULT 8  
US-09-124-900-3  
; Sequence 3, Application US/09124900

Patent No. 6268484  
GENERAL INFORMATION:  
APPLICANT: KATINGER, Hermann  
APPLICANT: BUCHACHER, Andrea  
APPLICANT: ERNST, Wolfgang  
APPLICANT: BALLAUN, Claudia  
APPLICANT: PURTSCHER, Martin  
APPLICANT: TRKOLA, Alexandra  
APPLICANT: PREDL, Renate  
APPLICANT: SCHMATZ, Christine  
APPLICANT: KLIMA, Annelies  
APPLICANT: STEINDL, Franz  
APPLICANT: MÜSTER, Thomas  
TITLE OF INVENTION: HIV-Vaccines  
FILE REFERENCE: 1939-112P  
CURRENT APPLICATION NUMBER: US/09/124,900  
CURRENT FILING DATE: 1998-07-30  
PRIOR APPLICATION NUMBER: PCT/EP95/01481  
PRIOR FILING DATE: 1995-04-19  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 3  
LENGTH: 1015  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
-09-124-900-3  
Query Match 45.6%; Score 61.5; DB 3; Length 1015;  
Best Local Similarity 78.6%; Pred. No. 3.1;  
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
1 KETWTWTWTEWSQ 14  
||| |||||:  
562 KET-WETWTWTEYWQ 574  
SULT 9  
-08-463-028-9  
Sequence 9, Application US/08463028  
Patent No. 6610476  
GENERAL INFORMATION:  
APPLICANT: CHANG, Nancy T.  
APPLICANT: GALLO, Robert C.  
APPLICANT: WONG-STAAI, Flossie  
TITLE OF INVENTION: CLONING AND EXPRESSION OF HTLV-III DNA  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Morgan & Finnegan, L.L.P.  
STREET: 345 Park Avenue  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10154-0053  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,028  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/693,866  
FILING DATE: 23-JAN-1985  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/659,339  
FILING DATE: 10-OCT-1984  
ATTORNEY/AGENT INFORMATION:  
NAME: Serunian, Leslie A.  
REGISTRATION NUMBER: 35,353  
REFERENCE/DOCKET NUMBER: 2026-4193U3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 758-4800  
TELEFAX: (212) 751-6849  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1015 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE: HTLV-III  
ORGANISM: HTLV-III  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..1015  
OTHER INFORMATION: /note= "pol protein of HTLV-III"  
US-08-463-028-9  
Query Match 45.6%; Score 61.5; DB 4; Length 1015;  
Best Local Similarity 78.6%; Pred. No. 3.1;  
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
QY 1 KETWTWTWTEWSQ 14  
||| |||||:  
Db 562 KET-WETWTWTEYWQ 574  
RESULT 10  
US-07-743-357-2  
Sequence 2, Application US/07743357  
Patent No. 5858646  
GENERAL INFORMATION:  
APPLICANT: Kang, Yong C.  
TITLE OF INVENTION: Polypeptide having immunological  
TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: KIRBY EADES GALE BAKER  
STREET: Box 3432, Station D  
CITY: Ottawa  
STATE: Ontario  
COUNTRY: Canada  
ZIP: K1M 1H8  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/743,357  
FILING DATE: 21-AUG-1991  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/CA90/00062  
FILING DATE: 23-FEB-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Gale, Edwin J.  
REGISTRATION NUMBER: 28,584  
REFERENCE/DOCKET NUMBER: 30924-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (613) 237-6900  
TELEFAX: (613) 237-0045  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1016 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Human immunodeficiency virus type 1

STRAIN: BH102  
07-743-357-2

Query Match 45.6%; Score 61.5; DB 2; Length 1016;  
Best Local Similarity 78.6%; Pred. No. 3.1;  
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

1 KETWETWTEWSQ 14  
|||||:|  
562 KET-WETWTEWQ 574

ULT 11

07-743-357-3  
sequence 3, Application US/07743357  
atent No. 5858646

GENERAL INFORMATION:

APPLICANT: Kang, Yong C.

TITLE OF INVENTION: Polypeptide having immunological

TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: KIRBY EADES GALE BAKER

STREET: Box 3432, Station D

CITY: Ottawa

STATE: Ontario

COUNTRY: Canada

ZIP: K1M 1H8

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/743.357

FILING DATE: 21-AUG-1991

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/CA90/00062

FILING DATE: 23-FEB-1990

ATTORNEY/AGENT INFORMATION:

NAME: Gale, Edwin J.

REGISTRATION NUMBER: 28,584

REFERENCE/DOCKET NUMBER: 30924-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (613) 237-6900

TELEFAX: (613) 237-0045

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 1016 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: protein

HYPOTHETICAL: NO

FRAGMENT TYPE: internal

ORIGINAL SOURCE:

ORGANISM: Human immunodeficiency virus type 1

STRAIN: BH5

07-743-357-3

Query Match 45.6%; Score 61.5; DB 2; Length 1016;  
Best Local Similarity 78.6%; Pred. No. 3.1;  
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

1 KETWETWTEWSQ 14  
|||||:|  
562 KET-WETWTEWQ 574

ULT 12

07-743-357-4

sequence 4, Application US/07743357

Patent No. 5858646

GENERAL INFORMATION:

APPLICANT: Kang, Yong C.

TITLE OF INVENTION: Polypeptide having immunological

TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: KIRBY EADES GALE BAKER

STREET: Box 3432, Station D

CITY: Ottawa

STATE: Ontario

COUNTRY: Canada

ZIP: K1M 1H8

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/743.357

FILING DATE: 21-AUG-1991

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/CA90/00062

FILING DATE: 23-FEB-1990

ATTORNEY/AGENT INFORMATION:

NAME: Gale, Edwin J.

REGISTRATION NUMBER: 28,584

REFERENCE/DOCKET NUMBER: 30924-2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (613) 237-6900

TELEFAX: (613) 237-0045

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 1016 amino acids

TYPE: amino acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: protein

HYPOTHETICAL: NO

FRAGMENT TYPE: internal

ORIGINAL SOURCE:

ORGANISM: Human immunodeficiency virus type 1

STRAIN: PV22

US-07-743-357-4

Query Match 45.6%; Score 61.5; DB 2; Length 1016;  
Best Local Similarity 78.6%; Pred. No. 3.1;  
Matches 11; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

QY 1 KETWETWTEWSQ 14  
|||||:|  
Db 562 KET-WETWTEWQ 574

RESULT 13

US-07-743-357-5

Sequence 5, Application US/07743357

Patent No. 5858646

GENERAL INFORMATION:

APPLICANT: Kang, Yong C.

TITLE OF INVENTION: Polypeptide having immunological

TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: KIRBY EADES GALE BAKER

STREET: Box 3432, Station D

CITY: Ottawa

STATE: Ontario

COUNTRY: Canada

ZIP: K1M 1H8

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/743,357  
FILING DATE: 21-AUG-1991  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/CA90/00062  
FILING DATE: 23-FEB-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Gale, Edwin J.  
REGISTRATION NUMBER: 28,584  
REFERENCE/DOCKET NUMBER: 30924-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (613) 237-6900  
TELEFAX: (613) 237-0045  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1016 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Human immunodeficiency virus type 1  
STRAIN: BRU  
-07-743-357-5

Query Match 45.6%; Score 61.5; DB 2; Length 1016;  
Best Local Similarity 78.6%; Pred. No. 3.1;  
Matches 11; Conservative 1; Mismatches 1; Gaps 1;  
1 KETWETWTEWSQ 14  
||| |||||:  
562 KET-WETWTEYMQ 574

SULT 14  
-07-743-357-6  
Sequence 6, Application US/07743357  
Patent No. 5858646  
GENERAL INFORMATION:  
APPLICANT: Kang, Yong C.  
TITLE OF INVENTION: Polypeptide having immunological  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KIRBY EADES GALE BAKER  
STREET: Box 3432, Station D  
CITY: Ottawa  
STATE: Ontario  
COUNTRY: Canada  
ZIP: K1M 1H8  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/743,357  
FILING DATE: 21-AUG-1991  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/CA90/00062  
FILING DATE: 23-FEB-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Gale, Edwin J.  
REGISTRATION NUMBER: 28,584  
REFERENCE/DOCKET NUMBER: 30924-2  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (613) 237-6900  
TELEFAX: (613) 237-0045  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 913 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Human immunodeficiency virus type 1  
STRAIN: MN  
US-07-743-357-6

Query Match 44.8%; Score 60.5; DB 2; Length 913;  
Best Local Similarity 76.9%; Pred. No. 3.7;  
Matches 10; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Oy 1 KETWETWTEWS 13  
Db 459 KET-WETWTEYT 470

RESULT 15  
US-07-743-357-8  
Sequence 8, Application US/07743357  
Patent No. 5858646  
GENERAL INFORMATION:  
APPLICANT: Kang, Yong C.  
TITLE OF INVENTION: Polypeptide having immunological  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KIRBY EADES GALE BAKER  
STREET: Box 3432, Station D  
CITY: Ottawa  
STATE: Ontario  
COUNTRY: Canada  
ZIP: K1M 1H8  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/743,357  
FILING DATE: 21-AUG-1991  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/CA90/00062  
FILING DATE: 23-FEB-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Gale, Edwin J.  
REGISTRATION NUMBER: 28,584  
REFERENCE/DOCKET NUMBER: 30924-2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (613) 237-6900  
TELEFAX: (613) 237-0045  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1003 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
FRAGMENT TYPE: internal  
ORIGINAL SOURCE:  
ORGANISM: Human immunodeficiency virus type 1  
STRAIN: RF  
US-07-743-357-8

very Match 41.9%; Score 56.5; DB 2; Length 1003;  
est Local Similarity 71.4%; Pred. No. 13;  
atches 10; Conservative 1; Mismatches 2; Indels 1; Gaps 1;  
1 KETWWETWWTEWSQ 14  
|||||  
549 KET-WEAWEWEYQ 561

rch completed: February 25, 2004, 14:49:22  
time : 25 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

protein - protein search, using sw model

on: February 25, 2004, 14:48:19 ; Search time 34 Seconds  
(without alignments)  
130.418 Million cell updates/sec

le: US-09-915-914B-7

fect score: 135

quence: 1 KETWETWTEWSQPKKKRV 21

ring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

urched: 809742 segs, 211153259 residues

al number of hits satisfying chosen parameters: 809742

imum DB seq length: 0

imum DB seq length: 2000000000

it-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

abase :

- Published Applications AA:\*
- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
  - 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
  - 3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep:\*
  - 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
  - 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
  - 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
  - 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
  - 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
  - 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*
  - 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*
  - 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
  - 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
  - 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
  - 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
  - 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
  - 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
  - 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
  - 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

sult No.	Score	Query Match	Length	DB ID	Description
1	135	100.0	21	10	US-09-915-914B-7
2	135	100.0	21	10	US-09-915-914B-36
3	135	100.0	21	14	US-10-226-956-291
4	135	100.0	21	14	US-10-077-555-13
5	135	100.0	21	14	US-10-211-088-309
6	135	100.0	21	14	US-10-403-339-32
7	135	100.0	21	15	US-10-360-275-37
8	127	94.1	21	15	US-10-372-003A-30
9	124	91.9	21	10	US-09-915-914B-16
10	123	91.1	21	10	US-09-915-914B-10
11	120	88.9	20	10	US-09-915-914B-8
12	108	80.0	19	10	US-09-915-914B-12
13	106	78.5	20	10	US-09-915-914B-9
14	105.5	78.1	20	10	US-09-915-914B-37
15	102	75.6	19	10	US-09-915-914B-11

Sequence 43, Appl  
Sequence 32, Appl  
Sequence 271, Appl  
Sequence 15, Appl  
Sequence 300, Appl  
Sequence 42, Appl  
Sequence 38, Appl  
Sequence 3, Appl  
Sequence 35, Appl  
Sequence 2, Appl  
Sequence 4, Appl  
Sequence 5, Appl  
Sequence 41, Appl  
Sequence 17, Appl  
Sequence 6, Appl  
Sequence 29, Appl  
Sequence 1, Appl  
Sequence 87, Appl  
Sequence 12, Appl  
Sequence 14, Appl  
Sequence 14, Appl  
Sequence 10, Appl  
Sequence 2, Appl  
Sequence 4, Appl  
Sequence 88, Appl  
Sequence 6, Appl  
Sequence 8, Appl  
Sequence 118, Appl  
Sequence 17, Appl  
Sequence 35, Appl

ALIGNMENTS

RESULT 1

US-09-915-914B-7  
; Sequence 7, Application US/09915914B  
; Publication No. US20030119725A1  
; GENERAL INFORMATION:  
; APPLICANT: Davita, Gilles  
; APPLICANT: Fernandez, Joseph  
; APPLICANT: Heitz, Frederic  
; APPLICANT: Morris, May  
; APPLICANT: Nery, Jean  
; APPLICANT: Archdeacon, John  
; APPLICANT: Horndorp, Kyle  
; TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS  
; FILE REFERENCE: AM-00105.P.1.1  
; CURRENT APPLICATION NUMBER: US/09/915,914B  
; CURRENT FILING DATE: 2001-07-26  
; PRIOR APPLICATION NUMBER: US 60/221,932  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic sequence  
US-09-915-914B-7

Query Match 100.0%; Score 135; DB 10; Length 21;  
Best Local Similarity 100.0%; Pred. No. 4.6e-09;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KETWETWTEWSQPKKKRV 21

Db 1 KETWETWTEWSQPKKKRV 21

RESULT 2

```
1-09-915-914B-36
Sequence 36, Application US/09915914B
Publication No. US20030119725A1
GENERAL INFORMATION:
APPLICANT: Divita, Gilles
APPLICANT: Fernandez, Joseph
APPLICANT: Heltz, Frederic
APPLICANT: Morris, May
APPLICANT: Mery, Jean
APPLICANT: Archdeacon, John
APPLICANT: Horndorp, Kyle
TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
FILE REFERENCE: AM-00105.P.1.1
CURRENT APPLICATION NUMBER: US/09/915,914B
CURRENT FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,932
PRIOR FILING DATE: 2000-07-31
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.1
SEQ ID NO 36
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic sequence
1-09-915-914B-36

Query Match          100.0%; Score 135; DB 10; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.6e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      1 KETWETWWTWTSQPKKKRKV 21
      |||||
      1 KETWETWWTWTSQPKKKRKV 21

RESULT 3
1-10-226-956-291
Sequence 291, Application US/10226956
Publication No. US20030060399A1
GENERAL INFORMATION:
APPLICANT: Brophy, Colleen
APPLICANT: Komalavilas, Padmini
APPLICANT: Panitch, Alyssa
APPLICANT: Joshi, Lokesh
APPLICANT: Seal, Brandon L.
TITLE OF INVENTION: REAGENTS AND METHODS FOR SMOOTH MUSCLE THERAPIES
FILE REFERENCE: ASU-1061-US
CURRENT APPLICATION NUMBER: US/10/226,956
CURRENT FILING DATE: 2002-08-23
PRIOR APPLICATION NUMBER: 60/314,535
PRIOR FILING DATE: 2001-08-23
NUMBER OF SEQ ID NOS: 320
SOFTWARE: PatentIn version 3.1
SEQ ID NO 291
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic peptide
3-10-226-956-291

Query Match          100.0%; Score 135; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.6e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      1 KETWETWWTWTSQPKKKRKV 21
      |||||
      1 KETWETWWTWTSQPKKKRKV 21

RESULT 4
3-10-077-555-13
Sequence 13, Application US/10077555
Publication No. US20030077289A1
GENERAL INFORMATION:
APPLICANT: Wang, Rong-fu
TITLE OF INVENTION: Use of Cell-Penetrating Peptides to Generate Antitumor Immunity
FILE REFERENCE: P02373US1/10200806
CURRENT APPLICATION NUMBER: US/10/077,555
CURRENT FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: US 60/268,687
PRIOR FILING DATE: 2001-02-15
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn version 3.1
SEQ ID NO 13
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic Peptide
US-10-077-555-13

Query Match          100.0%; Score 135; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.6e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      1 KETWETWWTWTSQPKKKRKV 21
      |||||
      1 KETWETWWTWTSQPKKKRKV 21

RESULT 5
US-10-211-088-309
Sequence 309, Application US/10211088
Publication No. US20030104479A1
GENERAL INFORMATION:
APPLICANT: Bright, Gary R.
APPLICANT: Premkumar, D. David
APPLICANT: Chen, Yih-Tai
TITLE OF INVENTION: No. US20030104479A1el Fusion Proteins And Assays For Molecular Bir
FILE REFERENCE: 01-1022-US
CURRENT APPLICATION NUMBER: US/10/211,088
CURRENT FILING DATE: 2002-10-15
PRIOR APPLICATION NUMBER: 60/309,395
PRIOR FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: 60/341,589
PRIOR FILING DATE: 2001-12-13
NUMBER OF SEQ ID NOS: 366
SOFTWARE: PatentIn version 3.1
SEQ ID NO 309
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Protein-derived transport peptide
US-10-211-088-309

Query Match          100.0%; Score 135; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.6e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      1 KETWETWWTWTSQPKKKRKV 21
      |||||
      1 KETWETWWTWTSQPKKKRKV 21

RESULT 6
US-10-405-339-32
Sequence 32, Application US/10405339
Publication No. US20030190364A1
GENERAL INFORMATION:
APPLICANT: Panitch, Alyssa
APPLICANT: Seal, Brandon
TITLE OF INVENTION: Biological Affinity Based Delivery Systems
FILE REFERENCE: 9138-0079US
US-10-405-339-32

Query Match          100.0%; Score 135; DB 14; Length 21;
Best Local Similarity 100.0%; Pred. No. 4.6e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      1 KETWETWWTWTSQPKKKRKV 21
      |||||
      1 KETWETWWTWTSQPKKKRKV 21
```

CURRENT APPLICATION NUMBER: US/10/405,339  
 CURRENT FILING DATE: 2003-04-01  
 PRIOR APPLICATION NUMBER: US 60/369,568  
 PRIOR FILING DATE: 2002-04-01  
 NUMBER OF SEQ ID NOS: 60  
 SOFTWARE: PatentIn version 3.1  
 EQ ID NO 32  
 LENGTH: 21  
 TYPE: PRT  
 ORGANISM: Artificial sequence  
 FEATURE:  
 OTHER INFORMATION: Synthetic peptide  
 10-405-339-32

Query Match 100.0%; Score 135; DB 14; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-09;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KETWETWTEWSQPKKKRV 21  
 |||||  
 1 KETWETWTEWSQPKKKRV 21

MULT 7  
 10-360-275-37  
 Sequence 37, Application US/10360275  
 Publication No. US2004001484A1  
 GENERAL INFORMATION:

APPLICANT: Active Motif  
 APPLICANT: Blinov, Vladimir  
 APPLICANT: Fernandez, Joseph  
 APPLICANT: Archdeacon, Dorothy  
 APPLICANT: Archdeacon, John  
 APPLICANT: Choob, Mikhail

TITLE OF INVENTION: OLIGONUCLEOTIDE ANALOGUES AND METHODS OF USE FOR MODULATING GENE  
 FILE REFERENCE: AM102.P.1.1.1US  
 CURRENT APPLICATION NUMBER: US/10/360,275  
 CURRENT FILING DATE: 2003-02-07  
 PRIOR APPLICATION NUMBER: US 10/072,975  
 PRIOR FILING DATE: 2002-02-09  
 PRIOR APPLICATION NUMBER: US 09/805,296  
 PRIOR FILING DATE: 2001-03-13  
 PRIOR APPLICATION NUMBER: US 60/189,190  
 PRIOR FILING DATE: 2000-03-14  
 NUMBER OF SEQ ID NOS: 37  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 37  
 LENGTH: 21  
 TYPE: PRT  
 ORGANISM: Artificial sequence  
 FEATURE:  
 OTHER INFORMATION: Synthetic sequence  
 10-360-275-37

Query Match 100.0%; Score 135; DB 15; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 4.6e-09;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KETWETWTEWSQPKKKRV 21  
 |||||  
 1 KETWETWTEWSQPKKKRV 21

MULT 8  
 10-372-003A-30  
 Sequence 30, Application US/10372003A  
 Publication No. US20030215846A1  
 GENERAL INFORMATION:  
 APPLICANT: Watt, Paul  
 APPLICANT: Thomas, Wayne  
 APPLICANT: Hopkins, Richard  
 TITLE OF INVENTION: Methods of constructing and screening

TITLE OF INVENTION: diverse expression libraries  
 FILE REFERENCE: FERIC40.00ICPI  
 CURRENT APPLICATION NUMBER: US/10/372,003A  
 CURRENT FILING DATE: 2003-02-21  
 PRIOR APPLICATION NUMBER: US 09/568,229  
 PRIOR FILING DATE: 2000-05-05  
 PRIOR APPLICATION NUMBER: US 60/132,711  
 PRIOR FILING DATE: 1999-05-05  
 NUMBER OF SEQ ID NOS: 81  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 30  
 LENGTH: 21  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: synthetic Pept1 peptide  
 US-10-372-003A-30

Query Match 94.1%; Score 127; DB 15; Length 21;  
 Best Local Similarity 95.2%; Pred. No. 3.7e-08;  
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KETWETWTEWSQPKKKRV 21  
 |||||  
 DB 1 KETWETWTEWSQPKKKRV 21

RESULT 9

US-09-915-914B-16  
 Sequence 16, Application US/09915914B  
 Publication No. US20030119725A1  
 GENERAL INFORMATION:  
 APPLICANT: Divita, Gilles  
 APPLICANT: Fernandez, Joseph  
 APPLICANT: Heitz, Frederic  
 APPLICANT: Morris, May  
 APPLICANT: Mery, Jean  
 APPLICANT: Archdeacon, John  
 APPLICANT: Hornsdorf, Kyle  
 TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS  
 FILE REFERENCE: AM-00105.P.1.1  
 CURRENT APPLICATION NUMBER: US/09/915,914B  
 CURRENT FILING DATE: 2001-07-26  
 PRIOR APPLICATION NUMBER: US 60/221,932  
 PRIOR FILING DATE: 2000-07-31  
 NUMBER OF SEQ ID NOS: 43  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 16  
 LENGTH: 21  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Synthetic sequence  
 NAME/KEY: MISC FEATURE  
 LOCATION: (10)..(10)  
 OTHER INFORMATION: X can be any amino acid or no amino acid  
 NAME/KEY: MISC FEATURE  
 LOCATION: (11)..(11)  
 OTHER INFORMATION: X can be any amino acid or no amino acid  
 US-09-915-914B-16

Query Match 91.9%; Score 124; DB 10; Length 21;  
 Best Local Similarity 90.5%; Pred. No. 8.2e-08;  
 Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KETWETWTEWSQPKKKRV 21  
 |||||  
 DB 1 KETWETWTEWSQPKKKRV 21

RESULT 10

US-09-915-914B-10  
 Sequence 10, Application US/09915914B



```
Publication No. US20030119725A1
GENERAL INFORMATION:
APPLICANT: Divita, Gilles
APPLICANT: Fernandez, Joseph
APPLICANT: Heitz, Frederic
APPLICANT: Morris, May
APPLICANT: Mery, Jean
APPLICANT: Archdeacon, John
APPLICANT: Hornrdorp, Kyle
TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
FILE REFERENCE: AM-00105.P.1.1
CURRENT APPLICATION NUMBER: US/09/915,914B
CURRENT FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,932
PRIOR FILING DATE: 2000-07-31
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.1
SEQ ID NO 10
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic sequence
-09-915-914B-10

Query Match          91.1%; Score 123; DB 10; Length 21;
Best Local Similarity 90.5%; Pred. No. 1.1e-07;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

      1 KETWTWTWTSQPKKKRKV 21
      |||||
      1 KETWTWTWTSQPKKKRKV 21

SULT 11
-09-915-914B-8
Sequence 8, Application US/09915914B
Publication No. US20030119725A1
GENERAL INFORMATION:
APPLICANT: Divita, Gilles
APPLICANT: Fernandez, Joseph
APPLICANT: Heitz, Frederic
APPLICANT: Morris, May
APPLICANT: Mery, Jean
APPLICANT: Archdeacon, John
APPLICANT: Hornrdorp, Kyle
TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
FILE REFERENCE: AM-00105.P.1.1
CURRENT APPLICATION NUMBER: US/09/915,914B
CURRENT FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,932
PRIOR FILING DATE: 2000-07-31
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.1
SEQ ID NO 8
LENGTH: 20
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic sequence
-09-915-914B-8

Query Match          88.9%; Score 120; DB 10; Length 20;
Best Local Similarity 89.5%; Pred. No. 2.2e-07;
Matches 17; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

      1 KETWTWTWTSQPKKKR 19
      |||||
      1 KETWTWTWTSQPKKKR 19

RESULT 12
-09-915-914B-12
Sequence 12, Application US/09915914B
Publication No. US20030119725A1
GENERAL INFORMATION:
APPLICANT: Divita, Gilles
APPLICANT: Fernandez, Joseph
APPLICANT: Heitz, Frederic
APPLICANT: Morris, May
APPLICANT: Mery, Jean
APPLICANT: Archdeacon, John
APPLICANT: Hornrdorp, Kyle
TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
FILE REFERENCE: AM-00105.P.1.1
CURRENT APPLICATION NUMBER: US/09/915,914B
CURRENT FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,932
PRIOR FILING DATE: 2000-07-31
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic sequence
US-09-915-914B-12

Query Match          80.0%; Score 108; DB 10; Length 19;
Best Local Similarity 88.9%; Pred. No. 5e-06;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

      4 WNETWTWTSQPKKKRKV 21
      |||||
      2 WNETWTWTSQPKKKRKV 19

QY      4 WNETWTWTSQPKKKRKV 21
DB      2 WNETWTWTSQPKKKRKV 19

RESULT 13
US-09-915-914B-9
Sequence 9, Application US/09915914B
Publication No. US20030119725A1
GENERAL INFORMATION:
APPLICANT: Divita, Gilles
APPLICANT: Fernandez, Joseph
APPLICANT: Heitz, Frederic
APPLICANT: Morris, May
APPLICANT: Mery, Jean
APPLICANT: Archdeacon, John
APPLICANT: Hornrdorp, Kyle
TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
FILE REFERENCE: AM-00105.P.1.1
CURRENT APPLICATION NUMBER: US/09/915,914B
CURRENT FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,932
PRIOR FILING DATE: 2000-07-31
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 20
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic sequence
US-09-915-914B-9

Query Match          78.5%; Score 106; DB 10; Length 20;
Best Local Similarity 84.2%; Pred. No. 8.8e-06;
Matches 16; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

      1 KETWTWTWTSQPKKKR 19
      |||||
      1 KETWTWTWTSQPKKKR 19

DB      1 KETWTWTWTSQPKKKR 19
      1 KETWTWTWTSQPKKKR 19

RESULT 14
US-09-915-914B-11
Sequence 11, Application US/09915914B
Publication No. US20030119725A1
GENERAL INFORMATION:
APPLICANT: Divita, Gilles
APPLICANT: Fernandez, Joseph
APPLICANT: Heitz, Frederic
APPLICANT: Morris, May
APPLICANT: Mery, Jean
APPLICANT: Archdeacon, John
APPLICANT: Hornrdorp, Kyle
TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
FILE REFERENCE: AM-00105.P.1.1
CURRENT APPLICATION NUMBER: US/09/915,914B
CURRENT FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,932
PRIOR FILING DATE: 2000-07-31
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.1
SEQ ID NO 11
LENGTH: 21
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic sequence
-09-915-914B-11

Query Match          91.1%; Score 123; DB 10; Length 21;
Best Local Similarity 90.5%; Pred. No. 1.1e-07;
Matches 19; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

      1 KETWTWTWTSQPKKKRKV 21
      |||||
      1 KETWTWTWTSQPKKKRKV 21
```

Search completed: February 25, 2004, 14:53:45  
Job time : 36 secs

```
09-915-914B-37
Sequence 37, Application US/09915914B
Publication No. US20030119725A1
GENERAL INFORMATION:
APPLICANT: Divita, Gilles
APPLICANT: Fernandez, Joseph
APPLICANT: Heitz, Frederic
APPLICANT: Morris, May
APPLICANT: Mery, Jean
APPLICANT: Archdeacon, John
APPLICANT: Horndorp, Kyle
TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
FILE REFERENCE: AM-00105.P.1.1
CURRENT APPLICATION NUMBER: US/09/915,914B
CURRENT FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,932
PRIOR FILING DATE: 2000-07-31
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.1
SEQ ID NO 37
LENGTH: 20
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic sequence
-09-915-914B-37

Query Match      78.1%; Score 105.5; DB 10; Length 20;
Best Local Similarity 85.7%; Pred. No. 1e-05;
Matches 18; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

1 KETWETWETWTSQPKKKKV 21
|||||
1 KETWETWETWTSQ-KKKKV 20

MULT 15
-09-915-914B-11
Sequence 11, Application US/09915914B
Publication No. US20030119725A1
GENERAL INFORMATION:
APPLICANT: Divita, Gilles
APPLICANT: Fernandez, Joseph
APPLICANT: Heitz, Frederic
APPLICANT: Morris, May
APPLICANT: Mery, Jean
APPLICANT: Archdeacon, John
APPLICANT: Horndorp, Kyle
TITLE OF INVENTION: PEPTIDE-MEDIATED DELIVERY OF MOLECULES INTO CELLS
FILE REFERENCE: AM-00105.P.1.1
CURRENT APPLICATION NUMBER: US/09/915,914B
CURRENT FILING DATE: 2001-07-26
PRIOR APPLICATION NUMBER: US 60/221,932
PRIOR FILING DATE: 2000-07-31
NUMBER OF SEQ ID NOS: 43
SOFTWARE: PatentIn version 3.1
SEQ ID NO 11
LENGTH: 19
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic sequence
-09-915-914B-11

Query Match      75.6%; Score 102; DB 10; Length 19;
Best Local Similarity 85.7%; Pred. No. 2.4e-05;
Matches 18; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

1 KETWETWETWTSQPKKKKV 21
|||||
1 KETWETW--TWSQPKKKKV 19
```